Beamforming Optimization for Two-Way Relay Channel

Haixia Chen

A dissertation submitted in fulfillment
of the requirements for the degree of
Doctor of Philosophy
of the
University of London.

Department of Electronic & Electrical Engineering
University College London

March 13, 2014

To my family.

Abstract

In this thesis, we focus on the optimization of the two-way relay channel (TWRC), which can double the data rate of communications comparing to the traditional one-way relay channel (OWRC). Because of the broadcasting nature of wireless transmissions, secure transmission is an appealing research topic. We take secrecy rate consideration into the optimization of the TWRC. Overall we provide near-optimal solutions for the secrecy rate maximization problems of the TWRC with imperfect channel state information (ICSI). A much lower complexity optimal SOCP solution is provided for SNR balancing of the TWRC without secrecy consideration. We first look at a flat fading TWRC network model with a multiple-input multiple-output (MIMO) relay where perfect channel state information (CSI) is assumed available. We then formulate an optimization problem, with the objective to minimize the relay's power usage under the constraints that the signal-to-noise ratio (SNR) of the two transceivers should exceed a preset threshold. A low-complexity optimal beamforming solution is provided to this optimization problem by reformulating it in the form of second-order cone programming (SOCP). Later in the thesis, we consider the presence of an eavesdropper and address the beamforming optimization for minimizing the relay's power with the constraints of the secrecy rates of the two transceivers. A semi-definite programming (SDP) based searching algorithm is proposed to find a near-optimal solution. For each search of the proposed approach, the previous non-convex optimization problem is transferred into an SDP problem, which can guarantee the optimality of the beamforming matrix. Afterwards, more realistic imperfect CSI (ICSI) situations are considered for the TWRC network models. As ICSI completely changes the structure and the property of the optimization problems, we reformulate the optimization problems into two scenarios. For the first case, we consider that the relay is an untrusted eavesdropper and in this case an SDP solution is provided to maximize the joint-decoding sum-secrecy rate. For the second case, we investigate the robust beamforming problems where the relay is trusted but there is an external eavesdropper, another SDP solution is provided to maximize the sum-secrecy rate.

Acknowledgements

I confirm that all the work here is my original work.

Contents

List of Figures

List of Tables

Notations / List of Abbreviations

- $\otimes$ denotes the Kronecker product
- $\odot$ denotes the Hadamard (elementwise) multiplication
- $(\cdot)^{\dagger}$ denotes the conjugate transpose of a matrix or vector
- $(\cdot)^{*}$ denotes the complex conjugate operation
- $(\cdot)^{T}$ denotes the transposition of a matrix or vector
- **1G** The First Generation
- **2G** The Second Genaration
- **3G** The Third Generation
- **4G** The Fourth Genaration
- **AF** Amplify and Forward
- **AWGN** Additive White Gaussian Noise
- **BC** Broadcast Channel
- **CF** Compress and Forward
- **CSI** Channel State Information
- **DF** Decode and Forward
- **ICSI** Imperfect Channel State Information
- **LMI** Linear Matrix Inequality
- **LOS** Line of Sight
- **MAC** Multiple Access Channel
- **MIMO** Multiple-input Multiple-output
- **MISO** Multiple-input Single-output
- **MISOME** Multiple-input Single-output Multiple-eavesdropper

- **MSE** Mean Square Error
- **NC** Network Coding
- **NC** Network Coding
- **OWRC** One-way Relay Channel
- **PNC** Physical layer network coding
- **QoS** Quality of Service
- **RF** Radio Frequency
- **SDP** Semi-definite programming
- **SIMO** Single-input Multi-output
- **SINR** Signal-tointerference-noise ratio
- **SNR** Signal-to-noise ratio
- **SOCP** Second-order Cone Program
- **SVD** Singular Value Decomposition
- **TWBF** Two-way Beamforming
- **TWRC** Two-way Relay Channel
- **TWZF** Two-way Zero-forcing Beamforming
- **WLAN** wireless local area networks
- **WMAN** Wireless Metropolitan Aera Network
- **WPAN** Wireless Personal Area Networks
- **WWAN** Wireless Wide Area Network
- **i.i.d** independent and identically distributed

Chapter 1

Introduction

1.1 Wireless Networks

Ever since the wireless telegraph was invented by Guglielmo Marconi [2] in 1896, wireless communications have dramatically changed the way people communicate. Compared to traditional wired networks, wireless networks provide flexibility and mobility, either to users or the service providers because some areas are difficult or too expensive to deploy network connection. In these cases, wireless communication is attractive. Besides, wireless communications provide the flexibility for quick and temporary network connections. For example, for exhibition or activity in an open area which has no network coverage, it would be cost ineffective to deploy cables to building a wired connection.

Also, the maintenance and upgrade of a wireless network is more convenient, time-effective and cost-effective. To upgrade the communication speed of wired network, sometimes new transmission hardware such as optical cable are required for replacement. It will take a long time to upgrade, and at the same time will cause in many cases interruptions to the existing services, and possibly bring inconvenience to the residences at that area. In contrast, the maintenance and management of wireless networks are much simpler. For this reason and with the fast developing wireless technologies, wireless networks are fast deployed to cover wide areas. Nowadays, information and communication can be provided conveniently into people's daily life. At the same time, people's requirement and expectation on wireless communications are increasing rapidly. Researchers and engineers are therefore working hard to improve the information rates given the precious spectrum and energy resources.

According to the differences from service range, mobility and transmission data rate, there are 4 different categories of wireless networks: wireless personal area network (WPAN), wireless local area network (WLAN), wireless metropolitan area network (WMAN) and wireless wide area network (WWAN).

1.1.1 WPAN

A WPAN is a personal area network connects devices within a short distance, usually several meters, of a person's walking range. It usually connects peripheral devices such as a printer, an IPAD or a mobile phone to a computer. The transmission rate of a WPAN connection varies from low to high, with different

technologies and standards. IEEE 802.15 is the standard for WPANs by IEEE Standards Association [7], including 7 task groups with transmission speed up to 3 Gbits/s. It uses Bluetooth (also known as IEEE 802.15.1) as the main technology, which provides low-rate transmission services with low-power consumption between devices. Another low-rate technology, ZigBee (IEEE 802.15.4), is intend to be less expensive and more simple than Bluetooth. WiMedia (IEEE 802.15.3) is the specification for high-rate applications of WPANs.

1.1.2 WLAN

A WLAN is a small local communication network connected by a series of devices and offering intra-network data exchange. The first workshop of WLANs was held in 1991by IEEE Standards Association [6] and had IEEE802.11 as its media access control and physical layer specification. Complex technical solutions, such as IEEE 802.11, 802.11a, 802.11b, 802.11g, 802.11n, and 802.11ac are applied to different geographic scopes of WLANs. Usually, a WLAN is spanned within one or several buildings owned by the same organisation. The transmission rate of a WLAN usually is much greater than a WPAN. Among all the WLAN standards, 802.11n offers high transmission speed at 100Mbits/s - 200Mbits/s, by using multiple-input multiple-output through (MIMO) spatial division multiplexing, and supports up to four spatial streams technique. Publication scheduled for early 2014, the IEEE 802.11ac is a WLAN standard providing high throughput on the 5 GHz band. It is expected that this specification will provide at least 1Gbits/s throughput of multi-stations, and at least 500Mbits/s throughput of a single link. MIMO (up to eight spatial streams) and beamforming functions are used by IEEE 802.11ac. Apple's new generation of airport is already supporting the 802.11ac.

1.1.3 WMAN

A WMAN is a comparably larger area of network than a WLAN by IEEE Standards Association [5] in 2002. A WMAN uses technologies different with WLANs and is specified under IEEE802.16 standards, offering high data rates (approximate 40Mbits/s of every channel) with guaranteed quality of service (QoS) to a potentially large customer base (up to tens of miles from the base station). There are two main drawbacks of the WMAN. One is that it is currently lack of mobility. The other one is its requirements of light of sight (LOS). If a customer does not have a clear LOS to the base station of the WMAN, the connection between the customer and the network can not be established.

1.1.4 WWAN

A WWAN network uses cellular network technologies to provide local, national and global voice and date services. The WWAN has been developed and deployed rapidly since 1980s. In 1981, the first generation (1G) mobile system was deployed. The second generation (2G) mobile system started to take place the 1G in 1992. Both the 1G and the 2G systems primarily focus on voice services. The third generation (3G) mobile system was developed and deployed in 2001. The focus of the 3G mobile system is on the data transmission rate, which is from 2Mbits/s and up to 14.4Mbits/s for stationary users. For high-speed moving users, the data transmission speed is up to 348Kbits/s. Since 2011, the fourth generation (4G) mobile system started to deployed globally, with a data transmission rate up to

1Gbps for stationary users and 100Mbits/s for high-speed mobile users.

1.2 Motivation and Objectives

Due to the harsh wireless signal's propagation environment which is constantly troubled by interference and fading, providing a stable high data rate for wireless services is a very challenging task. Alongside with already existed resource problems, such as the radio spectrum and the utilization of power, interference and fading are two fundamental problems raised from the broadcasting nature of wireless transmissions. These two factors make wireless transmission channels have an uncertainty issue.

We aim to optimize wireless networks, to properly design and deploy wireless networks which can maximize the usage of spectrum and power, and to reduce the influence of interferences and fadings. Of particular interest in recent years is the emerging two-way relay channel (TWRC) (first proposed by Zhang *et al.* [94]) which can extend wireless coverage using an intermediate relaying terminal as well as shorten the time by half for exchanging of messages between two terminals. Note that normally, for a half-duplex relay, it takes 4 time slots for two transceivers to exchange messages but with the two-way technologies, it is possible to take only 2 slots to accomplish this communication. An example of TWRC is shown in Fig.1.1

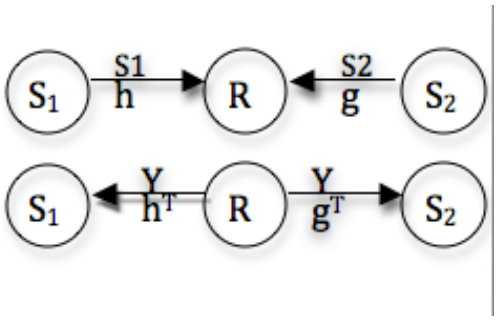

Figure 1.1: An example of TWRC.

The focus of this thesis is on the TWRC and in particular, our interest is on the use of a multiple-antenna (Roy *et al.* [66]) relaying terminal and we address several important optimization problems for the TWRC to enhance its reliable and secure information rate proposed at first by Shannon [71]. In 1988, Veen [81] took an overview of beamforming from the signal-processing's perspective. Beamforming is a linear signal processing technique used in sensor arrays for directional signal transmission or reception. It makes signals from particular angles have constructive interference and others experience destructive interference. It is also an approach of spatial filtering. The beamforming optimization for the two-way relay is not well known and this thesis has made three major contributions. In the first problem, our objective is to optimise the beamforming matrix at the relay for minimising the sum-power of the transceivers such that the received signal-to-noise ratios (SNRs, a measurement describing the level of designed signal to the level of background noise) of the transceivers exceed prescribed thresholds, given that the channel state information (CSI) (describes the signal's propagation from the transmitter

to the receiver) is perfectly known at the relay. Our contribution is the second-order cone programming (SOCP), a convex solution described by Boyd *et al.* [14] that finds the optimal beamforming matrix at the relay which is computationally with much lower complexity than the semidefinite programming (SDP, described by Boyd *et al.* [14]) solution in the literature. In addition, motivated by the great potential of providing security in the physical layer, widely known as physical layer security explained in details by Zhang *et al.* [1] (a scheme guarantees the message can not be decoded by illegal receivers through physical layer's transmission) , we also look at several optimization problems for securing the TWRC by optimising the beamforming matrix at the relay, with consideration of perfect CSI and imperfect CSI (ICSI) (only partial of the CSI is available, there is an uncertainty of the CSI) at the relay. In these cases, the sum-secrecy rate of the two transceivers is the performance metric. These optimizations permit us to keep the messages confidential from an eavesdropper by degrading the wiretap channel.

1.3 Overview of the Thesis and Contributions

The rest of the thesis is organised as follows. The introductory chapter is of Chapters 2 where the TWRC model and the researches to enhance secrecy during transmission at the physical layer, respectively, reviewed. The technical contributions of the thesis are in Chapters 3– 6.

In Chapter 3, our TWRC network model contains two single-antenna source nodes and a multiple-input multiple-output (MIMO) relay node. Perfect CSI is known at the relay node. This chapter's aim is to minimise the transmit power of the relay node while keeping the SNR at the end nodes higher than a preset threshold value. We propose an optimal low-complexity SOCP solution for the two-way relay beamforming. Computation complexity advantage of the SOCP solution is analysed.

Chapter 4 – Chapter 6 all address the TWRC with emphasis of enhancing not only the capacity but also the security. In Chapter 4, the TWRC in Chapter 3 is considered in the presence of a single-antenna passive eavesdropper. Again, perfect CSI is assumed for the optimization of the two-way beamforming matrix at the relay. Our problem of interest is to minimise the relay's transmit power subject to the secrecy rate constraints of the two transceivers. A 2D searching algorithm is proposed to find the near primal solution for this optimization problem. In each search, the previous non-convex optimization problem is being converted into an SDP problem, which can guarantee the optimality of the beamforming matrix.

ICSI situation, due to channel estimation errors or Doppler spread, is considered for the TWRC in Chapters 5 and 6. ICSI completely changes the structure and property of the optimal two-way relaying matrix, Chapter 6 reformulates the optimization problem to cope with bounded CSI errors that leads to the robust optimal solution in the presence of a passive external eavesdropper, while Chapter 6 deals with the same problem but the eavesdropper is the relay itself. Future work is given in the final chapter 7.

Our contributions have led to the following list of publications:

- H. Chen and K.-K. Wong, "Optimal two-way beamforming with perfect CSI: An SOCP formulation," in *London Commun. Sym. 2011*, UCL, London.

- L. Chen, K.-K. Wong, H. Chen, J. Liu, and G. Zheng, "Optimizing transmitter-receiver collaborative-relay beamforming with perfect CSI," *IEEE Commun. Letters*, vol. 15, no. 3, pp. 314–316, Mar. 2011.

- J. Ni, K.-K. Wong, Z. Fei, C. Xing, H. Chen, K.-F. Tong and J. Kuang, "Secrecy-rate balancing for two-user MISO interference channels," to appear in *IEEE Wireless Commun. Letters*, 2013.

- H. Chen and K.-K. Wong, "Optimal two-way beamforming with perfect CSI: An SOCP formulation," submitted to *Journal of Optimization*.

- H. Chen and K.-K. Wong, "Robust two-way relay-beamforming: An external eavesdropper case," submitted to *IEEE Trans. Signal Process.*.

- H. Chen and K.-K. Wong, "Robust two-way relay-beamforming: An untrusted relay case," to be submitted to *IEEE Trans. Wireless Commun.*.

Chapter 2

Background

This chapter reviews the basic background of wireless communications and gives a short introduction for the related concepts, such as TWRC, physical-layer security and beamforming optimization.

2.1 Wireless Communication System

Wireless communication corresponds to a message passing between a transmitter and a receiver over a wireless medium. Given a message for transmutation, the general function of a transmitter includes, but not limited to, (1) source coding (data compression) which maps the message from an information source to a sequence of alphabets (usually bits) such that the source symbols can be exactly recovered from the binary bits (lossless source coding) or recovered within some distortion (lossy source coding); (2) encryption which processes and encodes the alphabets to increase security and prevent it from illegitimate decoding; (3) channel coding which adds redundancy to the bit sequence for error protection due to noise, fading and interference; (4) modulation which converts the digital bit sequence into a proper waveform for efficient transmission, and (5) finally transmission from antenna. At the receiver side, it takes the radio signal and converts it back into electrical signal. Then the processing at the transmitter is reversed to recover the original message. Fig. 2.1 illustrates such a chain of processes.

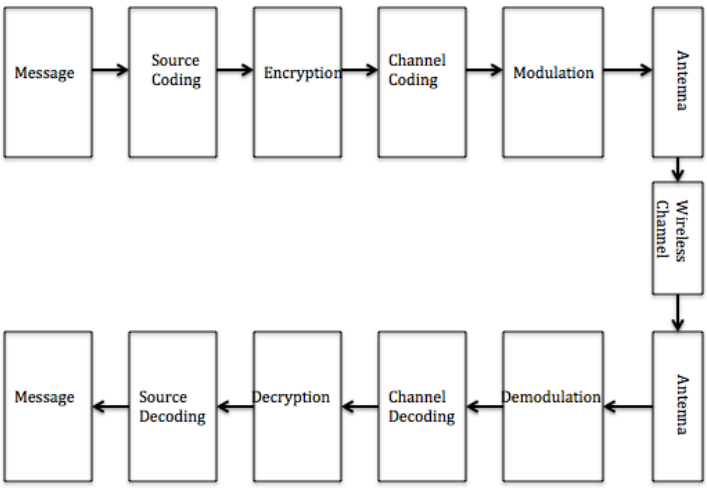

Figure 2.1: The processing of a typical wireless communication system.

In this thesis, we mainly focus on the influence of channel fading and additive noise in wireless communications system for optimization of wireless network (Power mimization, maximize the minimum SNR etc.). The above mentioned signal processing is presented in a simplified input-output relationship. Mathematically, it can be written as

$$y = hx + n, \tag{2.1}$$

where x is the transmit information-bearing digital symbol for the transmission antenna, h is the wireless channel coefficient, n is the additive white Gaussian noise with zero mean and power σ^2, and y is the received signal of the reciving antenna for further processing before decoding the original message. The channel characteristics is reviewed below.

2.1.1 Wireless fading channel

In wireless communication, the term "fading" refers to the deviation of attenuation of wireless signals after transmission, caused by multi-path propagation in the physical space. Precise mathematical expression of fading channel is unknown, as various environments (sometime intractable) make it too complex to describe with precision. Statistical models are therefore commonly used instead. There are currently various statistical models suitable for different propagation environments.

Fading can be classified either as fast and slow according to the speed of fluctuation compared to the speed of transmission, and can also be regarded as flat or selective in the frequency domain. In the work of this thesis, our emphasis is on slow frequency-flat fading channels because (1) *slow* fading is considered more challenging as channel coding would be ineffective without imposing severe delay, and (2) a frequency-selective fading channel can be easily converted into multiple parallel *frequency-flat* channels using technologies such as orthogonal frequency division multiplexing (OFDM).

The following models are commonly used for describing fading channel:

- **Rayleigh fading** occurs when the attenuation is a result of the superposition of infinitely many paths between the transmitter and the receiver, but none of the paths has a dominant signal, such as the signal via a line-of-sight (LOS) path. It is regarded as the worst-case scenario for wireless communications. This model also implies that the radio environment has a large number of scatterers. More accurately, Rayleigh fading also implies a complex Gaussian process where the amplitude is Rayleigh distributed but the phase is uniformly distributed. The probability density function (pdf) of the Rayleigh fading amplitude, α, is given by

$$p(\alpha) = \frac{2\alpha}{\Omega} \exp(-\frac{\alpha^2}{\Omega}), \text{ for } \alpha \geq 0, \tag{2.2}$$

  where $\Omega = \mathbb{E}(\alpha^2)$ is an important parameter determining the strength of the channel.

- **Rician fading** occurs when there is a dominant path such as LOS which is mixed with other indirect multiple paths. The parameter K is defined to characterise such channel:

$$K = \frac{\text{power in the dominant path}}{\text{power in the scattering paths}}. \tag{2.3}$$

When $K = 0$, the channel is reduced to a Rayleigh fading channel, while if $K = \infty$, the channel becomes a simple additive white Gaussian noise channel.

2.1.2 Performance Metrics

In this subsection, we review a number of popular metrics which will be useful in this thesis.

2.1.2.1 Capacity or Achievable Rate

Channel capacity is defined as the highest achievable rate that can be reliably decoded at the receiver over the channel. According to Shannon's Coding Theorem [72], if the information's transmission rate, denoted as R, is equal to or lower than the channel capacity C, then there exists scheme that can decode without any error the information. Conversely, if the information's transmission rate, R, is higher than the channel capacity C, the probability of decoding the information in error is close to 1.

For additive white Gaussian noise (AWGN) channels, channel capacity can be expressed as

$$C_{\mathsf{AWGN}} = \log\left(1 + \frac{P|h|^2}{\sigma^2}\right), \tag{2.4}$$

where P is the transmission power of the signal from the sender, σ^2 is the noise power and the channel fading coefficient h is considered fixed during the entire Gaussian noise process.

2.1.2.2 SNR

The received SNR, denoted as γ, is a useful metric that quantified the quality of the received signal. In (2.1), the SNR is simply given as

$$\gamma = \frac{\text{Power of the desired signal at the output}}{\text{Power of the noise at the output}} = \frac{P|h|^2}{\sigma^2}. \tag{2.5}$$

Sometimes, it is measured in decibels, or dB. That is,

$$\mathrm{SNR}_{\mathsf{dB}} = 10\log_{10}\gamma \text{ (in dB)}. \tag{2.6}$$

More precisely, γ is regarded as the instantaneous SNR which depends on the channel realisation h and varies if the channel changes. Therefore, it is also useful to define the average SNR, $\bar{\gamma}$, which is averaged over the fading channels, i.e.,

$$\bar{\gamma} = \int_0^\infty \gamma p(|h|) d|h|. \tag{2.7}$$

2.1.2.3 Signal-to-Interference Plus Noise Ratio (SINR)

In wireless communication systems, co-channel interferences generally exist. For example, in a wireless ad-hop network, a node can receive signals simultaneously from its neighbours. Therefore, besides the desired signal, there are interference signals as well. Consider the received signal at a node as

$$y = hx_1 + gx_2 + \eta, \tag{2.8}$$

in which x_1 is the desired signal, x_2 is the interference signal and η is the noise. Then the SINR, denoted as β, can be defined as

$$\beta = \frac{\text{Power of the desired signal at the output}}{\text{Power of the interference signal and noise at the output}} = \frac{P_1|h|^2}{P_2|g|^2 + \sigma^2}, \tag{2.9}$$

where P_1 is the transmission power of the desired signal, P_2 is the transmission power of the interference signal, and σ^2 is the noise power at the receiver side. Similar to SNR, SINR can be averaged over the channel fading coefficients, h and g, to provide more meaningful reference in practical systems.

2.1.2.4 Outage Probability

Maintaining a high level of quality-of-service (QoS) is always the aim of wireless network. One useful metric, which is increasingly popular due to its accurate reflection to user experience, is outage probability. Within this, various outage events can be defined appropriately and popular examples include rate outage, SNR outage and so on. As an example, let us consider the outage rate probability. All other outage probability of other metrics can be defined similarly.

Given a threshold value for the transmission rate, r_0, the outage probability can be found by

$$\text{Outage Probability} = \mathsf{Prob}(\text{Achievable Rate} < r_0). \tag{2.10}$$

2.1.2.5 Mean-Square-Error or MSE

MSE is another important metric which directly gives a measure for the accuracy of estimation. For instance, if we denote s_1 as the message sending out by the source node, and $\hat{s1}$ is the estimated message derived from the destination node. Then the MSE is obtained as

$$\text{MSE} = \mathbb{E}[|s_1 - \hat{s_1}|^2],$$

where $\mathbb{E}[\cdot]$ returns the expectation of an input random entity. Minimising the MSE, widely known as MMSE, is one of the most sought criteria for designing communications systems.

2.2 Cooperative Relay Network

2.2.1 Network Coding (NC)

In traditional computer networks, switching is functioned by nodes relaying their received messages to one or multiple links. When a switching node works as an encoder (this function can achieved through a powerful processing function of the antennas in Fig. 2.1) , it is known that network capacity can be significantly improved Ahlswede *et al.* [4]. To further increase the transmission rate, Ahlswede *et al.* [4] proposed the concept of NC (network coding). According to Shannon, the maximum channel capacity can be achieved through the "max-flow min-cut" theorem by Elias *et al.* [26]. However, with the traditional switching and routing scheme, information can only be stored and forwarded separately without any functional process at the medium access control (MAC) layer, but with the proposed NC method, network nodes are designed to code the incoming data with appropriate coding methods, such

as exclusive-or (XOR) operation, linear operation by Li *et al.* [75] and so on, so that the maximum transmission capacity defined by the "max-flow min-cut" theorem of Shannon can be achieved.

In [38], an algebraic framework of NC using discrete random process was proposed and it was shown that sink nodes can reconstruct the original data from the source accurately as long as the system transfer matrix has full rank. NC can take place either at upper layers by Katti *et al.* [13, 74], or at the physical layer by Zhang *et al.* [94]. This thesis falls into the latter case where a network of signals are processed in the physical layer.

2.2.2 Relaying Technologies

Distance and interference set limitation for wireless transmission, and extending the range or coverage of wireless communications has been an important problem since the signal's power decreases polynomially as the distance goes. There is always a case that the receiver is outside the transmitter's reliable transmission range. However, with the broadcast nature of wireless communications, it is possible that nodes close to the transmitter or the receiver can help relay the message that is not intended for them. This is usually referred to as a cooperative relay network [68, 69] where in the half-duplex setting, the relay node processes and forwards the message-bearing signal it received at the previous time slot to enhance the signal quality of others. The processing at the relay can vary.

- **Amplify-and-forward (AF)**—The relay amplifies the received signal and sends it. For AF, there is generally no requirement of the relay to know the CSI. It is therefore comparatively a straightforward relaying scheme. Statistical properties of AF relay fading channels were studied by Patel *et al.* [63], and by Chen *et al.* [16] studied the block-fading multiple-access channel using AF relaying.

- **Compress-and-forward (CF)**—In CF, the relay quantises the received message and then forwards the quantised signal. As compression is needed, it requires higher computation than AF and at the destination side, additional information may be needed to uncompress the signal.

- **Decode-and-forward (DF)**—The relay first decodes the received message at first and then re-encode it before forwarding. In the case, the CSI of the source-relay channel is expected to be available at the relay so that it can decode the message properly before re-encoding it. Gomadam *et al.* [32], SNR maximisation was considered to determine the optimal relaying function. Secrecy is a setback for DF relaying. As the relay node is not the intended destination, its capability of decoding the message facilitates overhearing of the message. Thus, if using DF, some other schemes such as jamming should be provided to overcome the possibility of leaking the message.

In this thesis, motivated by the complexity advantage, we will only consider the use of AF relaying.

In wireless communications, user cooperation can provide diversity using efficient protocols such as Laneman *et al.* [49]. The cooperative strategies and capacity theorems for relaying networks were studied by Kramer *et al.* [29]. Detailed performance analysis of the AF, CF and DF relaying schemes was presented by Cui *et al.* [24]. Generally, there are two types of relaying channels: one way relay channel (OWRC) and TWRC (two-way relay channel).

Consider a simple cooperative relay network model with 3 nodes. Nodes S_1 and S_2 are source nodes, and node R is the relaying node. In TWRCs, the two transceivers send messages to each other through a relay node or multiple relays. In this thesis, we focus on the case where there is only one relay but the relay is equipped with multiple antennas. In this 3-node network, depending on how the message is exchanged in time, there are typically 3 types of transmission schemes, as depicted in Fig. 2.2.

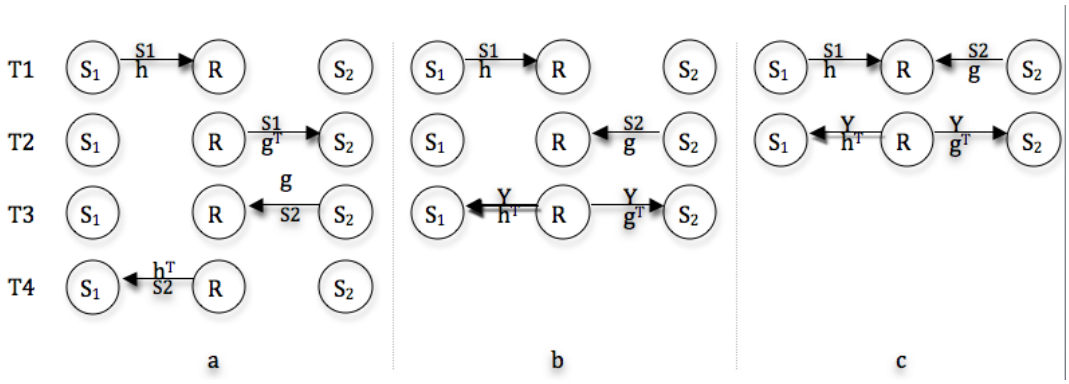

Figure 2.2: Various schemes for the cooperative relay channels.

2.2.3 The System Model for OWRC

The OWRC is a straightforward relaying scheme as illustrated in Fig. 2.2. If time division duplex (TDD) transmission is used, then 4 time slots will be required for the two nodes to exchange a message with each other. In Step 1, S_1 transmits message s_1 to the relay in the first time slot. Then R receives it as

$$r_1 = \mathbf{h}s_1 + \eta_{r_1}, \tag{2.11}$$

where $\mathbf{h}$ is a vector channel if R has multiple antennas, and η_{r_1} denotes the received noise. In Step 2, R processes the received signals with a matrix $\mathbf{W}$ (performs beamforming function of the relay by multiplying the received signal with matrix $\mathbf{W}$) and transmits it to S_2. Then S_2 receives the processed signal as

$$y_2 = \mathbf{g}^T\mathbf{W}r_1 + \eta_{s_2}, \tag{2.12}$$

where $\mathbf{g}$ is the vector channel between the relay and node S_2. After that, S_2 sends its message like S_1 just did, to yield the received signal at R:

$$r_2 = \mathbf{g}s_2 + \eta_{r_2}, \tag{2.13}$$

where η_{r_2} denotes the corresponding noise at the relay. Also, R will process the received vector signal by the same or different matrix $\mathbf{W}'$ and send it to S_1 over the channel $\mathbf{h}^T$, i.e.,

$$y_1 = \mathbf{h}^T\mathbf{W}r_2 + \eta_{s_1}. \tag{2.14}$$

This is the scheme (a) in Fig. 2.2.

The scheme (b) in Fig. 2.2 is another form of OWRC which however can shorten the time required to accomplish the message exchange. In particular, at the first 2 time slots, S_1 and S_2 transmit messages s_1 and s_2 to the relay separately. Then R receives the signals y_{r_1} and y_{r_2}, respectively, as

$$y_{r1} = \mathbf{h}s_1 + \eta_{r_1}, \tag{2.15}$$

$$y_{r2} = \mathbf{g}s_2 + \eta_{r_2}. \tag{2.16}$$

After receiving y_{r1} and y_{r2}, R then processes the two signals jointly with a beamforming matrix $\mathbf{W}$, and broadcasts the processed signal to both S_1 and S_2 which then receive:

$$y_1 = \mathbf{h}^T\mathbf{W}\mathbf{h}s_1 + \mathbf{h}^T\mathbf{W}\mathbf{g}s_2 + \mathbf{h}^T\mathbf{W}\eta_r + \eta_{s1}, \tag{2.17}$$

$$y_2 = \mathbf{g}^T\mathbf{W}\mathbf{h}s_1 + \mathbf{g}^T\mathbf{W}\mathbf{g}s_2 + \mathbf{g}^T\mathbf{W}\eta_r + \eta_{s2}. \tag{2.18}$$

The optimization problem for OWRC has been well studied. For example, [27] provided a convex optimization solution for the relay power minimization, with multiple relays and SNR constraints of a number of source-destination OWRC pairs.

2.2.4 The System Model for TWRC

Traditionally, a node cannot receive 2 different messages at the same time, and it is regarded as corruption of messages. However, Yuen *et al.* [90] presented the idea of physical layer NC and TWRC which uses this corruption of message instead of avoiding it, this resolves the problem and a node can receive different messages at the same time . With the mobile stations as the end nodes, and the base station as the relay node, In 2006, Larsson *et al.* [50] gave examples of the detailed transmission process of such TWRC and showed the gain improvement of the TWRC. Later in 2007, Kim *et al.* [47] introduced protocols for bidirectional relaying.

TWRC is a class of bidirectional channels that has received enormous attention recently, due to its high spectral efficiency of exchanging information between two terminals with the aid of an intermediate relaying terminal. The scheme (c) in Fig. 2.2 is a simple model of TWRC. It takes only 2 time slots for the two transceivers to exchange message with each other.

- Step 1: The two transceivers S_1 and S_2 transmit their messages simultaneously to R;
- Step 2: The relay R processes the received signal and broadcasts it in the next time slot.

For the scheme (c) in Fig. 2.2, there are 2 time slots required for the exchange of information, and the transmission rate of of S_1 and S_2 are:

$$R_1 = \frac{1}{2}\log\left(1 + \frac{P_2|\mathbf{h}^T\mathbf{W}\mathbf{g}|^2}{\|\eta_r\|^2\left(\|\mathbf{h}^T\mathbf{W}\|^2 + 1\right)}\right), \tag{2.19}$$

$$R_2 = \frac{1}{2}\log\left(1 + \frac{P_1|\mathbf{g}^T\mathbf{W}\mathbf{h}|^2}{\|\eta_r\|^2\left(\|\mathbf{g}^T\mathbf{W}\|^2 + 1\right)}\right). \tag{2.20}$$

The detailed capacity region of TWRC for various scenarios was discussed in Rankov *et al.* [65, 78, 82,

86].

Veen *et al.* [81] introduced beamformimg which is a linear signal processing function of the antenna (adjust amplitude or /and phase of the signals) for directional single transmission or reception and is attractive for improving the performance of wireless communications systems. By beamfrorming optimization in this thesis we mean choosing proper beamforming variables to satisfy some preset network matrics. Beamforming optimization gained huge attention recently from wireless networks ' research area. For example, downlink beamforming optimization problems were studied by Chalise *et al.* and Bengtsson *et al.* [10, 15]. Beamforming at the relay has also attracted much attention in recent years and has been addressed for the TWRC, which is the aim of this thesis. In particular, Zhang *et al.* [93] provided a thorough design for two-way beamforming for the case where a single multi-antenna relaying terminal forwards the array of the received noisy signals from the senders to the destination terminals in the AF fashion. Also, Havary-Nassab *et al.* [36] provided optimal cooperative beamfoming solutions for the TWRC with a number of single-antenna relays. It is worth pointing out here that there is fundamental difference between the literature and the work in this thesis. While Chapter 3 addresses the same optimization problem as by Zhang *et al.* [93], our contribution is to provide a much less complex optimal solution. On the other hand, Chapters 4–6 investigate the beamforming optimization problems for the TWRC with a single multi-antenna relay for enhancing the secrecy rates (to be introduced in Section 2.3) where physical layer security is considered at the same time. This formulation is novel and the proposed solutions are all not known before. We note that there has been early work for maximising secrecy rates for the TWRC using relay selection very recently by Cui *et al.* [23], but beamforming at a multiple-antenna relay was not considered in their work.

2.2.5 Robust Beamforming

Beamforming-based optimization is based on the assumption that exact knowledge of the CSI is available for the design. However, CSI is deemed to be imperfect due to estimation errors and exact CSI will be unavailable. In this case, the performance of the beamforming techniques designed for perfect CSI but operated using ICSI will be severely degraded. For this reason, robust optimization is required. Two robust approaches are commonly used. Ben-Tal *et al.* [9] studied these two approaches in details. In the first approach, the perturbation is assumed bounded with a maximum norm, and the constraints must be satisfied with the largest possible error, which leads to the worst-case optimization. The other approach considers the case in which the error is unbounded but statistically understood, which leads to robustness in the form of confidence level measured by probability. As a related work, the influence of CSI for TWRC was studied by Hammerstrom *et al.* [35]. Now, let us review the literature of using perfect CSI or ICSI for the TWRC.

2.2.5.1 TWRC with Perfect CSI

In TWRC, one problem of interest is to maximize the minimum of the SNRs from the two transceivers. Under the assumption of perfect CSI, for the network model that all the nodes are with multiple antennas, Wang *et al.* [85] provided a generalized fractional programming solution with constrained relay power. For the network models that the two transceivers nodes are with single antenna and the relay node is with

multiple antennas, Zhang *et al.* [93] provided an SDP solution to maximize the minimum of the SNRs. Besides, for the sum-rate maximization of the MIMO relay, Zhang *et al.* [93] provided a sub-optimal solution and Khabbazibasmenj *et al.* [3] provided a polynomial time solution. The optimal beamforming solution for sum-rate optimization of the MIMO relay is still an unsolved problem.

For the case of multiple single-antenna relays, Zeng *et al.* [59] presented the optimal weighted beamforming solution for maximising the sum-rate. Havary *et al.* [36] provided optimal cooperative beamfoming solutions (an optimal power minimization solution under the SNR constraint of each transceiver and an optimal SNR balancing solution) for the TWRC. A closed form solution for the relay power minimization with the SNR constraint to each transceiver is provided by Shahbazpanahi *et al.* [70] for the TWRC. For the second broadcasting stage of the TWRC with a single MIMO relay, a closed form sum-MSE minimization solution is provided by Kron *et al.* [48]. A selection of a relay out from K relays to achieve the best required performance metric is studied by Bletsas*et al.* [12] with an optimal solution provided to minimize the outage probability.

For the case that there are multiple-pair of AF TWRCs in the network, the sum rate maximization of the multi-pair TWRCs network with a total power constraint is studied by Zhang *et al.* [92]. A monotonic solution is provided to the optimization problem and a polyblock approximation algorithm for obtaining the global optimum is proposed. Wang *et al.* [83] considered the network optimization of K users each equipped with a single antenna and communicating via a relay with multi antennas. An unified approach is provided for the weighted sum MSE minimization and the sum-rate maximization with the constrained relay power.

Form the information theory side, the detailed capacity region of the TWRC in various scenarios was discussed by Rankov *et al.* [65, 78, 82, 86]. Achievable rate regions for the TWRC are presented by Ranhov *et al.* [65] with different cooperation schemes such as DF. Nam *et al.* [86] studied the capacity region especially for the Gaussian TWRC and an achievable scheme is provided. An unbound of the capacity for the TWRC is given by Zhang *et al.* [78]. Heath *et al.* [82] proved the linear scaling behaviour of the MIMO two-way relay channel with the increasing number of relay nodes.

2.2.5.2 TWRC with ICSI

The robust-optimal beamforming solution for collaborative-relays with imperfect but bounded CSI was investigated by Zheng *et al.* [34]. While using the MMSE as criterion, Xing *et al.* [25] provided an iterative algorithm (where each step is a quadratic matrix programming problem) to minimize the MMSE. In addition, Ubaidulla *et al.* [80], by modelling the CSI errors within an ellipsoidal region, a robust design which ensures the SNR constraint was presented. Recently, Aziz *et al.* [8] provided a robust beamforming approach under ICSI for a two-way relay system with analog NC. By using the S-procedure, the original constraints with infinite dimensions were converted into several linear matrix inequalities (LMIs) and the relaxed SDP was solved by applying rank-one relaxation. A principle eigenvector based rank-one reconstruction approach was proposed to reconstruct the solution. To reduce the outage probability, a hybrid approach that incorporates the non-robust approach when robust problem formulation is infeasible, was also presented.

On the other hand, Gharavol *et al.* [30] provided an iterative algorithm to minimize the sum MSE for joint optimization of the source precoders, the relay beamformer, and the destination equalisers for a two-way MIMO relay network with imperfect but bounded CSI. Moreover, Jun *et al.* [96] considered a stochastic model of the channel uncertainties in non-generative MIMO two-way relay network, and provided two algorithms for the robust joint source and relay optimization problem based on the MMSE criterion. A sub-optimal solution was also provided Tao *et al.* [56] for maximising the minimum worst-case SNR of the two sources subject to a total relay power budget. Ubaidulla *et al.* [80], with ICSI being bounded in an ellipsoidal area, a robust design which ensures that the SNR constraint is satisfied was presented. In addition, *et al.* Gharavol *et al.* [31] provided a set of LMIs that can be solved iteratively for the joint optimization of the source precoder, the relay beamformer and the destination equaliser in a non-regenerative relay network with norm bound CSI error matrix.

2.3 Physical Layer Security

For a message's transmission from the sender to the receiver(s), there is always possibility that the message can be overheard by any nodes (eavesdroppers) besides the legitimate receiver, as shown in Fig. 2.3. Therefore, security is always an issue, especially for wireless, where the signals are exposed.

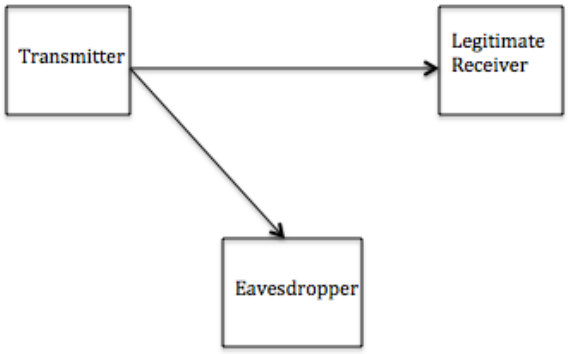

Figure 2.3: A basic model of a wiretap channel.

Prefect secrecy, which was first defined by Shannon [71], is said to be achieved when the transmitter delivers a positive information rate to the legitimate receiver that the eavesdropper cannot reveal any information. The first information-theoretic definition of secrecy rate of a wiretap channel was defined by Wyner [87], by measuring the conditional entropy through the equivocation-rate which is defined as

$$\frac{1}{n}H(W_1|Z_n), \tag{2.21}$$

where H stands for the entropy (see definition of information theory in more details in [21]).

Due to the broadcast nature in wireless communications, nodes within the transmission distance can receive the information-bearing signals. The nodes other than the desired receiver that can overhear the

message, are called eavesdroppers. In certain cases, the messages are required to be private and not be ciphered by others. Traditional ways to guarantee secrecy are by using cryptographic methods. A secret key is added to the original message and then sent out. The length of the key varies, but it requires either a comparable high level of computation power to decode the crypto-graphed message, or a key to decode it. The key is usually sent separately to the receiver, and thus once received the message, the legitimate receiver can decrypt it properly. This method is usually based on the upper layer's manipulation and not related to physical layer transmission. However, recent researches focus on manipulating the physical signals for providing an additional layer of security by degrading the signal quality of the eavesdropper.

Physical-layer security was first introduced by Wyner in the 70s [87]. With a degraded wiretap channel, he defined the notion of secrecy capacity to measure the maximum transmission rate from the source to the legitimate destination while the information leakage to the eavesdroppers is negligible. In 1978, Csiszar *et al.* [22] studied the secrecy capacity region for the discrete memoryless broadcast channel (BCC) with confidential messages. At the same year, Leung *et al.* [51] tudied the secrecy capacity region for the Gaussian wiretap channel. For BCC, subsequent researches in Liang and Liu [53, 54] presented the secrecy capacity region in Gaussian BCC and fading BCC, and exploiting fading to achieve secrecy. The secrecy region for other channels, such as multi-access channels, interference channels and relay channels has been widely studied in various papers by Liang *et al.* [33, 45, 46, 52, 53, 55, 60, 62], and more progress is being made. The secrecy capacity for single-input multiple-output (SIMO) slow fading channels is studied in Parada *et al.* [62]. The secrecy rate region for generalised MAC channel is studied by Liang *et al.* [52]. For a secure fading BCC channel with perfect CSI , the optimization problem about the minimal power that can minimize the outage probability, is derived by Liang *et al.* [53]. In the presence of an eavesdropper, the capacity region of the ergodic fading channel is explored with the consideration of both perfect CSI and ICSI cases by Gopala *et al.* [33]. The secrecy region of the Gaussian MIMO wiretap channel is studied by Khisti *et al.* [46], The role of multiple antennas for the secure transmission is studied by Khisti *et al.* [45], for the multiple-input single-output multiple-eavesdropper (MISOME) wiretap channel with both the sender and the eavesdropper have multiple antennas.

For the optimization of multiple-input single-output (MISO) channels using ICSI, Huang *et al.* [40] provided an SDP solution for maximising the MISO channel's secrecy rate, which is mathematically the difference of two log of the generalised Rayleigh quotient problems, and this type of problems can be solved by bisection searching methods. Beamforming strategies by Jing *et al.* [42] were also used to improve secrecy for various wiretap channels. In Mukherjee *et al.* [58], the authors discussed beamforming strategies for the MIMO wiretap channel. While for OWRC, Zhang *et al.* [43, 91] tackled the secrecy rate optimization using a null-space beamforming method.

2.3.1 Physical Layer Security of TWRC

The secrecy rate results mentioned above have in recent years attracted much attention in redefining wireless systems optimization, such as network secrecy sum-rate maximisation for TWRCs in Mukherjee *et al.* [57] and secrecy rate maximisation for OWRCs in Gan *et al.* [95]. From the information theory

point of view, coding theorems for perfect secrecy of TWRC were discussed by Schnurr *et al.* [67] with a result of an achievable region of half-duplex TWRC with CF relay nodes. However, achievable rate region for the TWRC under secrecy criterion is still an open problem. In Pierrot *et al.* [64], the rate region for strong secrecy with cooperative jamming was derived. He *et al.* [37] looked into the relationship between the feedback and the cooperative jamming of the TWRC and concluded that there was no difference between these two cases except the case that the channel performed the feedback functionality itself. According to their observation, feedback can increase the secrecy rate of the TWRC.

For the two-way relay AF MIMO relaying network, sum-rate maximization problems with or without secrecy consideration belong to the difference-of-convex (DC) programming problem. In Agis *et al.* [61], the authors analysed the sum-rate problem without secrecy consideration and provided two algorithms that can solve the problem in polynomial time. In addition, Mo *et al.* [41] derived the optimal structure to jointly optimize the source and relay beamformers in the two-phase TWRC with an untrusted relay. When the relay is unreliable, secure TWRC transmission was discussed in Luo *et al.* [20].

Friendly jammer discussed in Zhang *et al.* [1] is also a method to improve the secrecy rate of TWRC. The probability of successful detection for the system experiencing exponential path loss of TWRC was used as a security constraint in Fu *et al.* [28], and the optimization of minimizing the relay power based on this constraint was solved by searching methods. Chen *et al.* [18] proposed a cooperative jamming algorithm to select two or three intermediate nodes to enhance the security against the malicious eavesdropper, for a TWRC network model with a number of intermediate nodes and an eavesdropper. A two-phase distributed beamforming scheme is provided by Wang *et al.* [84] to maximize the secrecy rate for the TWRC with the existence of an extra eavesdropper. Secret key agreement schemes of the AF TWRC with the existence of artificial noises is studied by Shimizu *et al.* [73].

Chapter 3

SNR Maximisation for TWRCs

In this chapter, we investigate the optimisation of the relay beamforming for TWRCs. We use relay beamforming to achieve SNR balancing of the two source nodes of the TWRC network, assuming that perfect CSI is available for the optimisation. The aim is to minimise the relay transmission power, subject to the SNR constraints of the two transceivers. We prove that the objective function corresponding to this minimisation problem has a unique minimum and that the problem can be solved using SOCP.

Beamforming optimisation has been studied in the literature. For example, [36] considered the network model where all the nodes have only single antenna but there are multiple relays working together to form a two-way signal beam. In the case of multiple single-antenna relays, the relay beamforming matrix is a diagonal matrix. A semi-closed-form solution for this type of network was given in Shahbazpanahi *et al.* [70]. More relevant to this thesis, however, is the work by Zhang *et al.* [93] which provided a thorough design for two-way beamforming for the TWRC in which a single multi-antenna relaying node operating in an AF fashion was used. Their solution is globally optimal but relies on SDP with rank relaxation. The only drawback for their solution is that the computational complexity is still considerable.

The main contribution of this chapter is a SOCP formulation of the optimisation problem in Zhang *et al.* [93], which we show can be used to obtain the exact optimal two-way beamforming solution but at much less complexity than the SDP method. Such SOCP approach for the TWRC was not known before.

3.1 Problem Formulation

Consider the TWRC operating in TDD mode as shown in Fig. 3.1 in which we have two transceivers, labeled as S_1 and S_2, communicating with each other via an M-antenna relaying terminal, labeled as R. It is assumed that there is no direct link between S_1 and S_2. Denote the vector channels from S_1 and S_2 to R, respectively, as $\mathbf{h}$ and $\mathbf{g}$, which represent the flat-fading complex channel coefficients. Communications for the TWRC is achieved by two non-overlapping time slots.

During the first time slot, both S_1 and S_2 simultaneously transmit their messages to R. The signals received at R can be represented in vector form as

$$\mathbf{x} = \sqrt{P_1}\mathbf{h}s_1 + \sqrt{P_2}\mathbf{g}s_2 + \mathbf{v}, \tag{3.1}$$

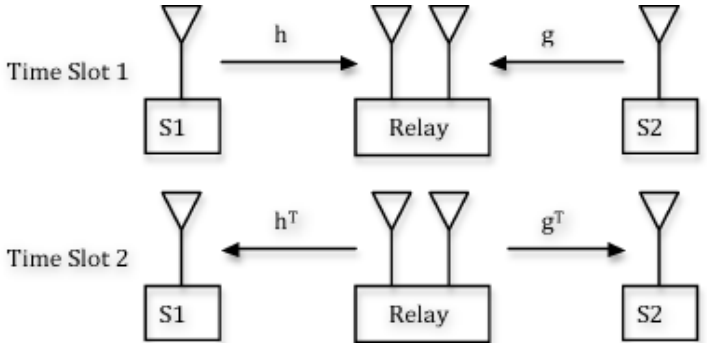

Figure 3.1: The TWRC Time-Slot Transmission.

where P_1 and P_2 are the respective transmit power of S_1 and S_2, s_1 and s_2 denote the symbols transmitted by S_1 and S_2, respectively, and $\mathbf{v} \in \mathcal{C}^M$ is the complex noise vector at R with independent and identically distributed (i.i.d.) zero-mean entries and $\mathbb{E}[\mathbf{v}\mathbf{v}^\dagger] = \sigma^2\mathbf{I}$ where $(\cdot)^\dagger$ denotes the complex conjugate transposition. At the second time slot, R transforms $\mathbf{x}$ by a complex matrix $\mathbf{W} \in \mathcal{C}^{M\times M}$ to give $\mathbf{W}\mathbf{x}$ and forwards it back to S_1 and S_2. As such, the signals received at S_1 and S_2 are given by

$$\begin{aligned} y_1 &= \mathbf{h}^T\mathbf{W}\mathbf{x} + \eta_1 \\ &= P_1\mathbf{h}^T\mathbf{W}\mathbf{h}s_1 + P_2\mathbf{h}^T\mathbf{W}\mathbf{g}s_2 + \mathbf{h}^T\mathbf{W}\mathbf{v} + \eta_1, \end{aligned} \tag{3.2}$$

$$\begin{aligned} y_2 &= \mathbf{g}^T\mathbf{W}\mathbf{x} + \eta_2 \\ &= P_1\mathbf{g}^T\mathbf{W}\mathbf{h}s_1 + P_2\mathbf{g}^T\mathbf{W}\mathbf{g}s_2 + \mathbf{g}^T\mathbf{W}\mathbf{v} + \eta_2, \end{aligned} \tag{3.3}$$

in which η_1 and η_2 denote the respective noise at S_1 and S_2 and they are assumed to be i.i.d. with zero mean and variance of σ^2.

In TWRCs, s_2 is intended for S_1. As s_1 is known for S_1 and with perfect CSI, the term carrying s_1 can be removed from y_1 to give

$$\tilde{y}_1 = \sqrt{P_2}\left(\mathbf{h}^T\mathbf{W}\mathbf{g}s_2 + \mathbf{h}^T\mathbf{W}\mathbf{v}\right) + \eta_1. \tag{3.4}$$

Similarly for S_2, we have

$$\tilde{y}_2 = \sqrt{P_1}\left(\mathbf{g}^T\mathbf{W}\mathbf{h}s_2 + \mathbf{g}^T\mathbf{W}\mathbf{v}\right) + \eta_2. \tag{3.5}$$

As a result, the SNRs at the terminals S_1 and S_2 are given by

$$\text{SNR at } \mathsf{S}_1 \equiv \gamma_1 = \frac{P_2|\mathbf{h}^T\mathbf{W}\mathbf{g}|^2}{\sigma^2\left(\|\mathbf{h}^T\mathbf{W}\|^2 + 1\right)}, \tag{3.6}$$

$$\text{SNR at } \mathsf{S}_2 \equiv \gamma_2 = \frac{P_1|\mathbf{g}^T\mathbf{W}\mathbf{h}|^2}{\sigma^2\left(\|\mathbf{g}^T\mathbf{W}\|^2 + 1\right)}, \tag{3.7}$$

where $\|\cdot\|$ returns the Euclidean norm of a vector.

Our goal is to optimally obtain the beamforming weight coefficients as well as the transceiver's transmit powers, for minimising the total transmit relay power subject to the SNR constraints for the

transceivers. This consideration leads to the SNR balancing problem:

$$\begin{aligned} &\min_{\mathbf{W}} \; P_1\|\mathbf{W}\mathbf{h}\|^2 + P_2\|\mathbf{W}\mathbf{g}\|^2 + \mathsf{trace}(\mathbf{W}\mathbf{W}^\dagger)\sigma^2 \\ &\text{s.t.} \; \begin{cases} \gamma_1 \geq \Gamma_1, \\ \gamma_2 \geq \Gamma_2, \end{cases} \end{aligned} \tag{3.8}$$

where Γ_1 and Γ_2 are the respective target SNRs at S_1 and S_2.

3.2 An SOCP Formulation

In this section, we present a SOCP formulation to obtain the optimal solution to (3.8). To do so, firstly, we study the structure of the beamforming matrix $\mathbf{W}$. In our network model, $\mathbf{W}$ is only operated with the channel vectors $\mathbf{h}$ and $\mathbf{g}$. Therefore, we find it useful to define

$$\mathbf{A} = [\mathbf{h}\ \mathbf{g}] \in \mathbb{C}^{M\times 2} \tag{3.9}$$

and as in [93] we write the SVD of $\mathbf{A}$ as

$$\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger, \tag{3.10}$$

where $\mathbf{U} = [\mathbf{U}^{\|}\ \mathbf{U}^{\perp}] \in \mathbb{C}^{M\times M}$ is a unitary matrix with $\mathbf{U}^{\|} \in \mathbb{C}^{M\times 2}$ and $\mathbf{U}^{\perp} \in \mathbb{C}^{M\times(M-2)}$, $\mathbf{V} \in \mathbb{C}^{2\times 2}$ is another unitary matrix and

$$\mathbf{\Sigma} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \\ 0 & 0 \\ \vdots & \end{bmatrix} \in \mathbb{C}^{M\times 2}, \tag{3.11}$$

in which $\lambda_1 \geq \lambda_2 \geq 0$ are the singular values of $\mathbf{A}$. Since $\mathbf{U}$ is a unitary matrix, we have $\mathbf{U}^{-1} = \mathbf{U}^T$, and we can find a similar matrix of $\mathbf{W} = \mathbf{U}^*\mathbf{A}\mathbf{U}$. As a result, $\mathbf{W}$ can be expressed as [93]

$$\mathbf{W} = [(\mathbf{U}^{\|})^*\ (\mathbf{U}^{\perp})^*] \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{bmatrix} [\mathbf{U}^{\|}\ \mathbf{U}^{\perp}]^\dagger, \tag{3.12}$$

where $\mathbf{B} \in \mathbb{C}^{2\times 2}$, $\mathbf{C} \in \mathbb{C}^{2\times(M-2)}$, $\mathbf{D} \in \mathbb{C}^{(M-2)\times 2}$ and $\mathbf{E} \in \mathbb{C}^{(M-2)\times(M-2)}$.

Without loss of generality, from (3.12), $\mathbf{W}$ can be expressed as

$$\begin{aligned} \mathbf{W} &= [(\mathbf{U}^{\|})^*\ (\mathbf{U}^{\perp})^*] \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{bmatrix} [\mathbf{U}^{\|}\ \mathbf{U}^{\perp}]^\dagger \\ &= (\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger + (\mathbf{U}^{\|})^*\mathbf{C}(\mathbf{U}^{\perp})^\dagger + (\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\|})^\dagger + (\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^\dagger. \end{aligned} \tag{3.13}$$

First, from (3.13), it can be shown that

$$\begin{aligned}&|\mathbf{h}^T\mathbf{W}\mathbf{g}|^2\\ =&|\mathbf{h}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger\mathbf{g}+\mathbf{h}^T(\mathbf{U}^{\|})^*\mathbf{C}(\mathbf{U}^{\perp})^\dagger\mathbf{g}+\mathbf{h}^T(\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\|})^\dagger\mathbf{g}+\mathbf{h}^T(\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^\dagger\mathbf{g}|^2\\ =&|\mathbf{h}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^H\mathbf{g}|^2,\end{aligned} \tag{3.14}$$

and

$$\begin{aligned}&|\mathbf{g}^T\mathbf{W}\mathbf{h}|^2\\ =&|\mathbf{g}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger\mathbf{h}+\mathbf{g}^T(\mathbf{U}^{\|})^*\mathbf{C}(\mathbf{U}^{\perp})^\dagger\mathbf{h}+\mathbf{g}^T(\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\|})^\dagger\mathbf{h}+\mathbf{g}^T(\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^\dagger\mathbf{h}|^2\\ =&|\mathbf{g}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^H\mathbf{h}|^2.\end{aligned} \tag{3.15}$$

Furthermore, we also have

$$\begin{aligned}\|\mathbf{W}^H\mathbf{h}^*\|^2&=\|\mathbf{B}^H(\mathbf{U}^{\|})^T\mathbf{h}^*\|^2+\|\mathbf{C}^H(\mathbf{U}^{\|})^T\mathbf{h}^*\|^2+\|\mathbf{D}^H\mathbf{U}^{\perp}\mathbf{h}^*\|^2+\|\mathbf{D}^H\mathbf{U}^{\perp}\mathbf{h}^*\|^2\\ &=\|\mathbf{B}^H\mathbf{U}^T\mathbf{h}^*\|^2+\|\mathbf{C}^H\mathbf{U}^T\mathbf{h}^*\|^2,\end{aligned} \tag{3.16}$$

and

$$\begin{aligned}\|\mathbf{W}^H\mathbf{g}^*\|^2&=\|\mathbf{B}^H(\mathbf{U}^{\|})^T\mathbf{g}^*\|^2+\|\mathbf{C}^H(\mathbf{U}^{\|})^T\mathbf{g}^*\|^2+\|\mathbf{D}^H\mathbf{U}^{\perp}\mathbf{g}^*\|^2+\|\mathbf{D}^H\mathbf{U}^{\perp}\mathbf{g}^*\|^2\\ &=\|\mathbf{B}^H(\mathbf{U}^{\|})^T\mathbf{h}^*\|^2+\|\mathbf{C}^H(\mathbf{U}^{\|})^T\mathbf{h}^*\|^2.\end{aligned} \tag{3.17}$$

Since the vector of $\mathbf{U}^{\perp}$ is the null-space of $\mathbf{A}$, which is the vector span space of $\mathbf{h}$ and $\mathbf{g}$, we have

$$\mathbf{U}^{\perp}\mathbf{h}=\mathbf{U}^{\perp}\mathbf{g}=0. \tag{3.18}$$

Substituting this structure into the SNR constraints, it can be easily seen that γ_1 and γ_2 are not related to $\mathbf{D}$ and $\mathbf{E}$ and furthermore for minimising the relaying power, the matrices $\mathbf{C}$, $\mathbf{D}$ and $\mathbf{E}$ should all be set to zeros. Hence, we can write

$$\mathbf{W}=(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger. \tag{3.19}$$

Apparently, $\mathsf{rank}(\mathbf{W})=2$ or the optimal $\mathbf{W}$ lies on the vector span of $\mathbf{h}$and $\mathbf{g}$. Now, by stacking all the elements of $\mathbf{B}$ into a vector, we define

$$\mathbf{w}\triangleq\mathsf{vec}(\mathbf{B})\in\mathcal{C}^{4\times1}. \tag{3.20}$$

Also, $\mathbf{h}_1 \triangleq (\mathbf{U}^{\parallel})^{\dagger}\mathbf{h} \in \mathbb{C}^{2\times 1}$, $\mathbf{g}_1 \triangleq (\mathbf{U}^{\parallel})^{\dagger}\mathbf{g} \in \mathbb{C}^{2\times 1}$, and let

$$\mathbf{f}_1 \triangleq \mathsf{vec}(\mathbf{h}_1\mathbf{g}_1^T) \tag{3.21}$$

$$\mathbf{f}_2 \triangleq \mathsf{vec}(\mathbf{g}_1\mathbf{h}_1^T) \tag{3.22}$$

$$\mathbf{U}_1 \triangleq P_1\mathbf{h}_1\mathbf{h}_1^{\dagger} + P_2\mathbf{g}_1\mathbf{g}_1^{\dagger} + \sigma^2\mathbf{I} \tag{3.23}$$

$$\mathbf{U}_2 \triangleq [\mathsf{diag}(\mathbf{U}_1^T, \mathbf{U}_1^T)]^{\frac{1}{2}} \tag{3.24}$$

$$\mathbf{H} \triangleq \begin{bmatrix} [\mathbf{h}_1]_1 & 0 & [\mathbf{h}_1]_2 & 0 \\ 0 & [\mathbf{h}_1]_1 & 0 & [\mathbf{h}_1]_2 \end{bmatrix} \tag{3.25}$$

$$\mathbf{G} \triangleq \begin{bmatrix} [\mathbf{g}_1]_1 & 0 & [\mathbf{g}_1]_2 & 0 \\ 0 & [\mathbf{g}_1]_1 & 0 & [\mathbf{g}_1]_2, \end{bmatrix} \tag{3.26}$$

where the notation "$[\mathbf{a}]_n$" returns the nth entry of $\mathbf{a}$ (a similar notation is also used for denoting the entry of a matrix). Thus, (3.8) becomes

$$\mathbb{P} \mapsto \begin{cases} \min\limits_{\mathbf{w}} \ \|\mathbf{U}_2\mathbf{w}\|^2 \\ \text{s.t.} \ \begin{cases} P_2|\mathbf{f}_1^T\mathbf{w}|^2 \geq \Gamma_1\left[\sigma^2\left(\|\mathbf{H}\mathbf{w}\|^2+1\right)\right], \\ P_1|\mathbf{f}_2^T\mathbf{w}|^2 \geq \Gamma_2\left[\sigma^2\left(\|\mathbf{G}\mathbf{w}\|^2+1\right)\right]. \end{cases} \end{cases} \tag{3.27}$$

The rest of this section is devoted to show that (3.27) has a SOCP solution and such solution is optimal.

To do so, we consider the following SOCP problem:

$$\mathbb{P}_{\mathsf{SOCP}} \mapsto \begin{cases} \min\limits_{\mathbf{w}} \ \|\mathbf{U}_2\mathbf{w}\|^2 \\ \text{s.t.} \ \begin{cases} P_2(\mathbf{f}_1^T\mathbf{w})^2 \geq \Gamma_1\left[\sigma^2\left(\|\mathbf{H}\mathbf{w}\|^2+1\right)\right], \\ P_1(\mathbf{f}_2^T\mathbf{w})^2 \geq \Gamma_2\left[\sigma^2\left(\|\mathbf{G}\mathbf{w}\|^2+1\right)\right], \\ \mathsf{Im}(\mathbf{f}_1^T\mathbf{w}) = \mathsf{Im}(\mathbf{f}_2^T\mathbf{w}) = 0. \end{cases} \end{cases} \tag{3.28}$$

The additional constraints in (3.28) can be rewritten as

$$\frac{\mathsf{Re}([\mathbf{W}]_{1,2}) - \mathsf{Re}([\mathbf{W}]_{2,1})}{\mathsf{Im}([\mathbf{W}]_{1,2}) - \mathsf{Im}([\mathbf{W}]_{2,1})} = -\frac{\mathsf{Re}([\mathbf{h}_1]_2[\mathbf{g}_1]_1) - \mathsf{Re}([\mathbf{h}_1]_1[\mathbf{g}_1]_2)}{\mathsf{Im}([\mathbf{h}_1]_2[\mathbf{g}_1]_1) - \mathsf{Im}([\mathbf{h}_1]_1[\mathbf{g}_1]_2)} = a. \tag{3.29}$$

To facilitate our analysis for the SOCP problem, we re-express (3.28) into the form of real vectors and

matrices by defining

$$\tilde{\mathbf{w}} = [\mathsf{Re}(\mathbf{w})^T \ \mathsf{Im}(\mathbf{w})^T]^T, \tag{3.30}$$

$$\mathbf{F}_1 = \begin{bmatrix} \mathsf{Re}(\mathbf{f}_1) & -\mathsf{Im}(\mathbf{f}_1) \\ \mathsf{Im}(\mathbf{f}_1) & \mathsf{Re}(\mathbf{f}_1) \end{bmatrix}^T \begin{bmatrix} \mathsf{Re}(\mathbf{f}_1) & -\mathsf{Im}(\mathbf{f}_1) \\ \mathsf{Im}(\mathbf{f}_1) & \mathsf{Re}(\mathbf{f}_1) \end{bmatrix}, \tag{3.31}$$

$$\mathbf{F}_2 = \begin{bmatrix} \mathsf{Re}(\mathbf{f}_2) & -\mathsf{Im}(\mathbf{f}_2) \\ \mathsf{Im}(\mathbf{f}_2) & \mathsf{Re}(\mathbf{f}_2) \end{bmatrix}^T \begin{bmatrix} \mathsf{Re}(\mathbf{f}_2) & -\mathsf{Im}(\mathbf{f}_2) \\ \mathsf{Im}(\mathbf{f}_2) & \mathsf{Re}(\mathbf{f}_2) \end{bmatrix}, \tag{3.32}$$

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathsf{Re}(\mathbf{H}) & -\mathsf{Im}(\mathbf{H}) \\ \mathsf{Im}(\mathbf{H}) & \mathsf{Re}(\mathbf{H}) \end{bmatrix}^T \begin{bmatrix} \mathsf{Re}(\mathbf{H}) & -\mathsf{Im}(\mathbf{H}) \\ \mathsf{Im}(\mathbf{H}) & \mathsf{Re}(\mathbf{H}) \end{bmatrix}, \tag{3.33}$$

$$\tilde{\mathbf{G}} = \begin{bmatrix} \mathsf{Re}(\mathbf{G}) & -\mathsf{Im}(\mathbf{G}) \\ \mathsf{Im}(\mathbf{G}) & \mathsf{Re}(\mathbf{G}) \end{bmatrix}^T \begin{bmatrix} \mathsf{Re}(\mathbf{G}) & -\mathsf{Im}(\mathbf{G}) \\ \mathsf{Im}(\mathbf{G}) & \mathsf{Re}(\mathbf{G}) \end{bmatrix}, \tag{3.34}$$

$$\tilde{\mathbf{U}}_2 = \begin{bmatrix} \mathsf{Re}(\mathbf{U}_2) & -\mathsf{Im}(\mathbf{U}_2) \\ \mathsf{Im}(\mathbf{U}_2) & \mathsf{Re}(\mathbf{U}_2) \end{bmatrix}^T \begin{bmatrix} \mathsf{Re}(\mathbf{U}_2) & -\mathsf{Im}(\mathbf{U}_2) \\ \mathsf{Im}(\mathbf{U}_2) & \mathsf{Re}(\mathbf{U}_2) \end{bmatrix}. \tag{3.35}$$

As a result, $\mathbb{P}_{\mathsf{SOCP}}$ becomes

$$\mathbb{P}_{\mathsf{SOCP}} \mapsto \begin{cases} \min\limits_{\mathbf{w}} \ \tilde{\mathbf{w}}^T \tilde{\mathbf{U}}_2 \tilde{\mathbf{w}} \\ \text{s.t.} \begin{cases} \tilde{\mathbf{w}}^T \tilde{\mathbf{A}}_1 \tilde{\mathbf{w}} \geq \Gamma_1 \sigma^2, \\ \tilde{\mathbf{w}}^T \tilde{\mathbf{A}}_2 \tilde{\mathbf{w}} \geq \Gamma_2 \sigma^2, \\ \dfrac{[\tilde{\mathbf{w}}]_2 - [\tilde{\mathbf{w}}]_3}{[\tilde{\mathbf{w}}]_6 - [\tilde{\mathbf{w}}]_7} = a, \end{cases} \end{cases} \tag{3.36}$$

where $\tilde{\mathbf{A}}_1 = P_2 \mathbf{F}_1 - \Gamma_1 \sigma^2 \tilde{\mathbf{H}}$ and $\tilde{\mathbf{A}}_2 = P_1 \mathbf{F}_2 - \Gamma_2 \sigma^2 \tilde{\mathbf{G}}$.

Theorem 3.2.1. *The problem $\mathbb{P}$ in (3.8) has an SOCP optimal solution which is the optimal solution to $\mathbb{P}_{\mathsf{SOCP}}$ in (3.28) or (3.36).*

Proof. Given the optimal solution to $\mathbb{P}$, say $\mathbf{x}$, we create a vector, $\mathbf{y}$, which differs only on the second, third, sixth and seventh elements. We prove our result by showing that it is feasible to find a suitable $\mathbf{y}$ that satisfies the SOCP constraints in (3.36) and achieves the same minimum objective value, i.e.,

$$\begin{cases} \mathbf{d}^T \mathbf{U} \mathbf{d} + \mathbf{d}^T \mathbf{u} = 0, \\ \mathbf{d}^T \tilde{\mathbf{A}}_1^{\dagger} \mathbf{d} + \mathbf{d}^T \mathbf{a}_1 \geq 0, \\ \mathbf{d}^T \tilde{\mathbf{A}}_2^{\dagger} \mathbf{d} + \mathbf{d}^T \mathbf{a}_2 \geq 0, \\ [\mathbf{d}]_3 - [\mathbf{d}]_2 = a([\mathbf{d}]_6 - [\mathbf{d}]_7 + [\mathbf{x}]_7 - [\mathbf{x}]_6) + [\mathbf{x}]_2 - [\mathbf{x}]_3, \end{cases} \tag{3.37}$$

where

$$\mathbf{d} = \begin{bmatrix} y_2 - x_2 \\ y_3 - x_3 \\ y_6 - x_6 \\ y_7 - x_7 \end{bmatrix}, \tag{3.38}$$

$\mathbf{U}$ is the sub-matrix of $\tilde{\mathbf{U}_2}$ extracting the (i,j)th entries for $i,j = 2,3,6,7$, $\tilde{\mathbf{A}_1}^{\dagger}$ and $\tilde{\mathbf{A}}_2^{\dagger}$ are matrices defined in a similar way from $\tilde{\mathbf{A}_1}$ and $\tilde{\mathbf{A}_2}$, respectively. Likewise, $\mathbf{u}$ is defined as the column vector with the entry being the product of the ith row vector of $\tilde{\mathbf{U}_2}$ and $\mathbf{x}$ for $i = 2,3,6,7$, and the vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ are defined similarly with the entries of $\tilde{\mathbf{A}_1}$ and $\tilde{\mathbf{A}_2}$, respectively, instead of $\tilde{\mathbf{U}_2}$.

Apparently, there are four variables in $\mathbf{d}$, which are required to meet two equality and two inequality constraints, totally four independent conditions. Therefore, it is feasible to find the solutions for $\mathbf{d}$ to achieve this and as a result, there are SOCP solutions that can achieve the same minimum objective value for $\mathbb{P}$. Now, given the optimal solution to $\mathbb{P}$, say $\mathbf{x}$, we create a vector, $\mathbf{y}$,

$$\mathbf{y} = \mathbf{x} + \mathbf{y}' \tag{3.39}$$

where $[\mathbf{y}']_n = [\mathbf{y}]_n - [\mathbf{x}]_n$. We prove our result by showing that it is feasible to find a suitable $\mathbf{y}$ that satisfies the SOCP constraints in (3.36) and achieves the same minimum objective value, i.e.,

$$\mathbf{x}^T\tilde{\mathbf{U}_2}\mathbf{x} = \mathbf{y}^T\tilde{\mathbf{U}_2}\mathbf{y}, \tag{3.40}$$

$$\mathbf{y}^T\tilde{\mathbf{A}_1}\mathbf{y} \geq \mathbf{x}^T\tilde{\mathbf{A}_1}\mathbf{x} \geq \Gamma_1\sigma^2, \tag{3.41}$$

$$\mathbf{y}^T\tilde{\mathbf{A}_2}\mathbf{y} \geq \mathbf{x}^T\tilde{\mathbf{A}_2}\mathbf{x} \geq \Gamma_2\sigma^2, \tag{3.42}$$

$$\frac{[y]_2 - [y]_3}{[y]_6 - [y]_7} = a. \tag{3.43}$$

Apparently, (3.40) and (3.43) are necessary conditions for the optimality of the SOCP solution, whereas (3.41) and (3.42) are the sufficient conditions rather than necessary. Taking $[\mathbf{y}']_n = [\mathbf{y}]_n - [\mathbf{x}]_n$ into (3.40)–(3.42), we get

$$\left\|\tilde{\mathbf{U}}_2^{\frac{1}{2}}\mathbf{y}' + \tilde{\mathbf{U}}_2^{-\frac{1}{2}}\mathbf{c}_1\right\|^2 = \mathbf{c}_1^T\tilde{\mathbf{U}}_2^{-1}\mathbf{c}_1, \tag{3.44}$$

$$\left\|\tilde{\mathbf{A}}_1^{\frac{1}{2}}\mathbf{y}' + \tilde{\mathbf{A}}_1^{-\frac{1}{2}}\mathbf{a}_1\right\|^2 \geq \mathbf{a}_1^T\tilde{\mathbf{A}}_1^{-1}\mathbf{a}_1, \tag{3.45}$$

$$\left\|\tilde{\mathbf{A}}_2^{\frac{1}{2}}\mathbf{y}' + \tilde{\mathbf{A}}_2^{-\frac{1}{2}}\mathbf{a}_2\right\|^2 \geq \mathbf{a}_2^T\tilde{\mathbf{A}}_2^{-1}\mathbf{a}_2, \tag{3.46}$$

where $\mathbf{c}_1 = \tilde{\mathbf{U}}_2\mathbf{x}$, $\mathbf{a}_1 = \tilde{\mathbf{A}}_1\mathbf{x}$ and $\mathbf{a}_2 = \tilde{\mathbf{A}}_2\mathbf{x}$. The cone $\|\mathbf{x}\| \leq t$ with $\mathbf{x} \in \mathcal{R}^2$ is illustrated in Fig. 3.2.

It is apparent that $\|\tilde{\mathbf{U}}_2^{\frac{1}{2}}\mathbf{y}' + \tilde{\mathbf{U}}_2^{-\frac{1}{2}}\mathbf{c}_1\|^2 = \mathbf{c}_1^T\tilde{\mathbf{U}}_2^{-1}\mathbf{c}_1$, $\|\tilde{\mathbf{A}}_1^{\frac{1}{2}}\mathbf{y}' + \tilde{\mathbf{A}}_2^{-\frac{1}{2}}\mathbf{a}_1\|^2 \leq \mathbf{a}_1^T\tilde{\mathbf{A}}_1^{-1}\mathbf{a}_1$ and $\|\tilde{\mathbf{A}}_2^{\frac{1}{2}}\mathbf{y}' + \tilde{\mathbf{A}}_2^{-\frac{1}{2}}\mathbf{a}_2\|^2 \leq \mathbf{a}_2^T\tilde{\mathbf{A}}_2^{-1}\mathbf{a}_2$ are three cones originated from $\tilde{\mathbf{U}}_2^{-\frac{1}{2}}\mathbf{c}_1$, $\tilde{\mathbf{A}}_1^{-\frac{1}{2}}\mathbf{a}_1$ and $\tilde{\mathbf{A}}_2^{-\frac{1}{2}}\mathbf{a}_2$, respectively, and with height $\sqrt{\mathbf{c}_1^T\tilde{\mathbf{U}}_2^{-1}\mathbf{c}_1}$, $\sqrt{\mathbf{a}_1^T\tilde{\mathbf{A}}_1^{-1}\mathbf{a}_1}$ and $\sqrt{\mathbf{a}_2^T\tilde{\mathbf{A}}_2^{-1}\mathbf{a}_2}$, respectively, in the 3D case. From this, we can see the intersection (3.44)–(3.46). We can thus simply prove that the intersection of (3.44) and (3.48) is

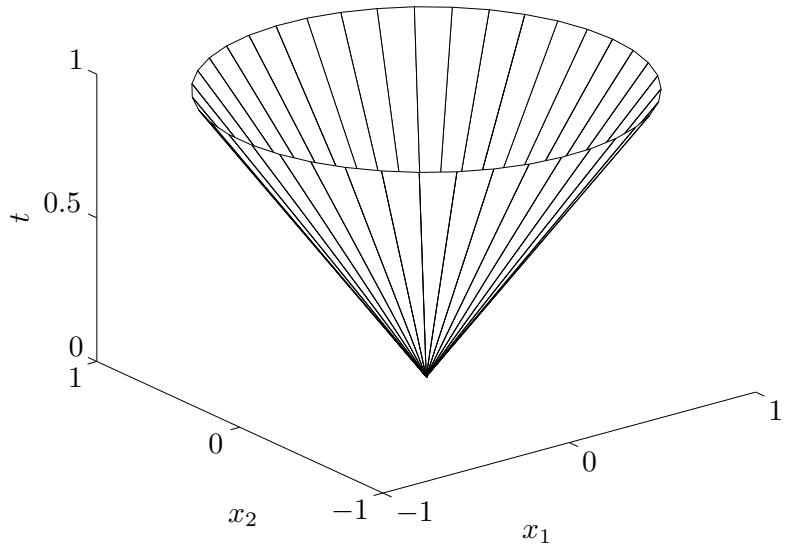

Figure 3.2: The SOC cone.

not empty:

$$\|\tilde{\mathbf{U}}_2^{\frac{1}{2}}\mathbf{y}' + \tilde{\mathbf{U}}_2^{-\frac{1}{2}}\mathbf{c}_1\|^2 = \mathbf{c}_1^T\tilde{\mathbf{U}}_2^{-1}\mathbf{c}_1, \tag{3.47}$$

$$\mathbf{q}\mathbf{y}' = b, \tag{3.48}$$

where $\tilde{\mathbf{U}}_2\mathbf{x} = \mathbf{c}_1$; $\mathbf{q} = [0, 1, -1, 0, 0, -a, a, 0]^T$, and $b = (x_3 - x_2) + a(x_6 - x_7)$.

Each of the constraints is a feasible problem: (3.44) is an SOC, which is $\in \mathcal{R}^n$ and the hyperplane (3.48) $\in \mathcal{R}^{n-1}$ is also a feasible problem. Now, rewrite $\mathbf{U}_2^{\frac{1}{2}} = [\mathbf{u}_1^T, \mathbf{u}_2^T, \mathbf{u}_8^T]^T$. Then we can recast (3.48) into $y_2' = b + y_3' + ay_6' - ay_7'$ and taking it into (3.44), we get

$$\left\|\tilde{\mathbf{U}}_3[y_1', y_3', y_4', y_5', y_6', y_7', y_8'] + (b\mathbf{u}_2^T + \tilde{\mathbf{U}}_2^{-\frac{1}{2}}\mathbf{c}_1)\right\|^2 = \mathbf{c}_1^T\tilde{\mathbf{U}}_2^{-1}\mathbf{c}_1, \tag{3.49}$$

where $\tilde{\mathbf{U}}_3 = [\mathbf{u}_1^T, \mathbf{u}_3^T, \mathbf{u}_4^T, \mathbf{u}_5^T, (a\mathbf{u}_2 + \mathbf{u}_6)^T, (\mathbf{u}_7 - a\mathbf{u}_2)^T, \mathbf{u}_8^T]^T$. As a result, (3.44) is still feasible, which completes the proof. The intersection is shown in Fig. 3.3. □

Note that Theorem 3.2.1 is unique for TWRCs with only one relaying terminal and if there are more than one relays, then the SOCP solution will no longer be optimal. This can be explained by adding one more inequality constraint similar to (3.41) and (3.42). In that case, we will have 3 hyperplane equality requirement $\mathbf{q}_i\mathbf{y}' = b_i$, for $i = 1, 2, 3$, the intersection of the 3 hyperplanes puts a special requirement of the channel coefficient vectors between the source nodes and the relay. Because of co-phasing, we have vectors $\mathbf{q}_1 = \mathbf{q}_2 = \mathbf{q}_3$. The intersectional area only exists if $b_1 = b_2 = b_3$; otherwise, the 3 hyperplanes will have to be parallel with each other and the intersection will be empty. Therefore, we can see that our SOCP solution is specially fit for two inequality requirements (3.41) and (3.42).

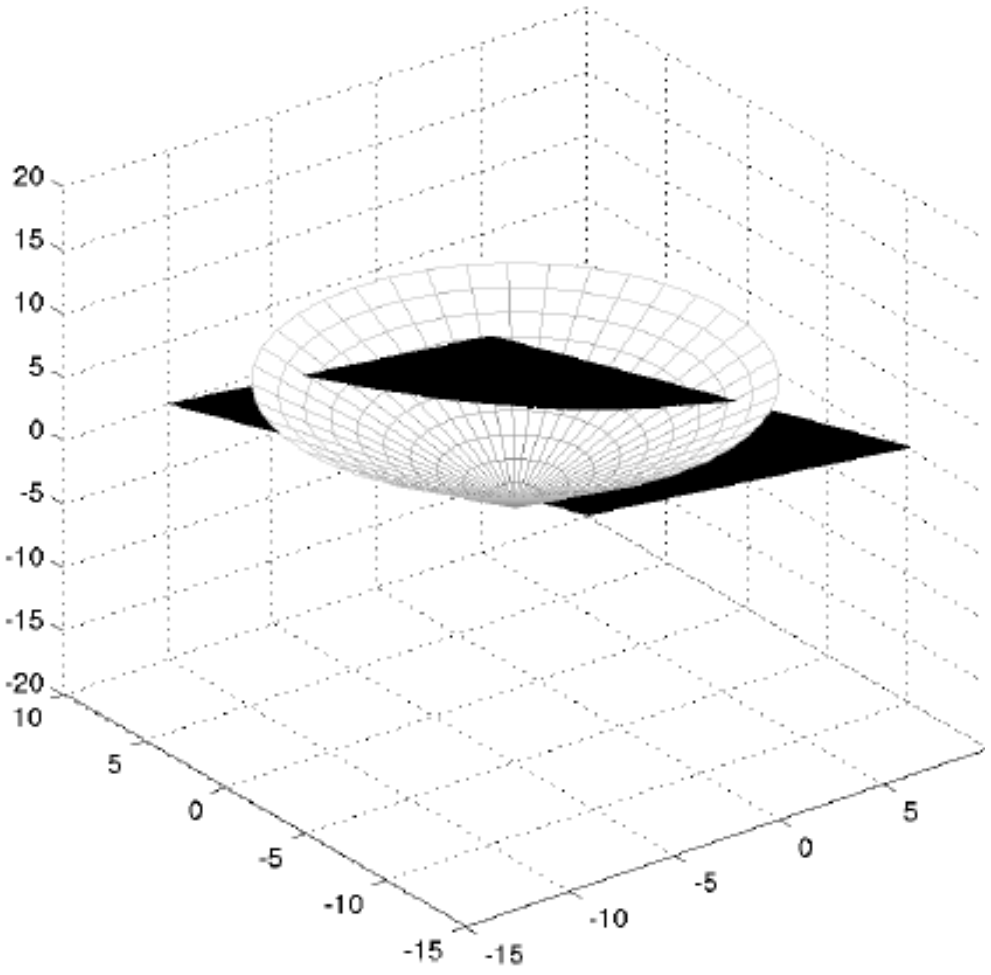

Figure 3.3: The intersection.

3.3 Complexity Analysis

For the SDP method with relaxiation in Zhang. [93], the worst case complexity is $O(mn^3 + m^2n^2)$ for computing $\mathsf{trace}(\mathbf{CX})$ [14] using general SDP solvers such as SEDUMI [76] and CVX, where $\mathbf{C}$ is an $n \times n$ symmetric matrix, and m is the number of the constraints. Generally, the worst case complexity for SDP and SOCP can both be regarded as $O(N^{3.5})$ with N being the number of variables involved. Note that Zhang [93] always has 8×8 symmetric matrices $\tilde{\mathbf{A}}_1^\dagger$ and $\tilde{\mathbf{A}}_2^\dagger$ so $N = 8$. However, in our SOCP optimisation, we always have $m = 2$ and two asymmetric 4×8 matrices in the form of

$$\begin{bmatrix} \mathsf{Re}(\mathbf{H}) & -\mathsf{Im}(\mathbf{H}) \\ \mathsf{Im}(\mathbf{H}) & \mathsf{Re}(\mathbf{H}) \end{bmatrix} \text{ and } \begin{bmatrix} \mathsf{Re}(\mathbf{G}) & -\mathsf{Im}(\mathbf{G}) \\ \mathsf{Im}(\mathbf{G}) & \mathsf{Re}(\mathbf{G}) \end{bmatrix}. \tag{3.50}$$

Methods	Number of Variables	Order	dim	blocks
SDP	N = 36	l = 11	dim=67	2
Methods	Complexity	nnz(A)	nnz(ADA)	nnz(L)
SDP	$O(36^{3.5})$	84	1296	666
Methods	No. of Variables	Order	dim	blocks
SOCP	N= 12	l = 13	dim=30	6
Methods	Complexity	nnz(A)	nnz(ADA)	nnz(L)
SOCP	$O(8^{3.5})$	96	118	65

Table 3.1: Complexity comparisons for SDP in [93] and the proposed SOCP using SEDUMI.

3.4 Simulation Results

Table 3.3 lists the simulation related parameters for SDP and SOCP using SEDUMI. The rank-one approaching algorithm in [93] has r parallel inequalities with one variable each and can be computed in $O(r)$. We simulated 500 times for each (m, R) pair for the SDP solution with the number of antennas $m = [2, 12]$, SNR threshold $R = [2, 20]$, and the probability of getting a rank-1 solution appears to be 0 for each pair of (m, R) and all the rank are 8. Based on these, the matrix decomposition complexity is $O(512)$ and the rank-one approaching complexity is $O(8)$. For [93], the SDP optimisation dominates the computation time and our SOCP optimisation solution is much faster than [93].

For $P_1 = P_2 = 10db$, we simulated 1000 independent channel realisations for each (m, R) pair for the SDP solution in Zhang *et al.* [93] with the number of antennas $m = [2, 12]$, SNR threshold $R = [1, 10]$, and the probability of getting rank-1 solution is provided in Table 3.4. In this table, $\overline{r}$ shows the average rank of the solution and it indicates that with the increment of the antenna number of the relay, it is highly likely that SDP gives a non rank-one solution. In contrast, as our SOCP can always get the rank-1 optimal solution, Zhang *et al.* [93] will suffer even higher complexity of $O(r^3)$ for additional matrix decomposition and the rank-one approaching with additional complexity of $O(r)$. Fig. 3.4 illustrates the average relay transmit power performance against the source SNR for various number of antennas at the relay.

R \ m	2	4	6	8	10	12
r=1(2db)	100%	87.3%	64.5%	89.7%	71.2%	55.7%
$\overline{r}(2db)$	1	1.5	2.6	1.6	2.9	5.7
r=1(4db)	82.3%	100%	62.3%	95.3%	90.8%	92.5%
$\overline{r}(4db)$	2.1	1	4	1.2	1.3	1.9
r=1(6db)	93.4%	51.2%	92.7%	50.3%	91.2%	33.7%
$\overline{r}(6db)$	1.2	2.5	2.6	5.8	1.7	6.2
r=1(8db)	-	100%	79.8%	72.7%	90.3%	81.2%
$\overline{r}(8db)$	-	1	3.4	4	1.6	2.6
r=1(10db)	-	74.3%	69.8%	97.3%	100%	77.5%
$\overline{r}(10db)$	-	2.8	3.2	1.1	1	3
r=1(12db)	-	87.7%	37.4%	62.7%	80.3%	72.7%
$\overline{r}(12db)$	-	4.5	2.6	2.9	1.7	2.9
r=1(14db)	-	94.2%	62.7%	100%	51.8%	52.7%
$\overline{r}(14db)$	-	1.2	4.8	1	2.9	3.4
r=1(16db)	-	-	100%	33.8%	80.3%	45.3%
$\overline{r}(16db)$	-	-	1	6.3	2.1	3.7
r=1(18db)	-	-	89.6%	72.3%	56.9%	75.80%
$\overline{r}(18db)$	-	-	1.4	2.3	2.9	2.2
r=1(20db)	-	-	67.5%	81.5%	78.3%	47.8%
$\overline{r}(20db)$	-	-	4.6	2.3	2.9	5.7

Table 3.2: SOCP and SDP Simulation rank-one solution comparation

3.5 Conclusion

This chapter has presented an SOCP method for determining the optimal relay beamforming matrix for the SNR balancing problem in the TWRC with the assumption that the perfect CSI is available to all the nodes within the network. The objective function of the SNR balancing problem is to minimize the relay power, based on the constraints that SNR of each terminal nodes are above a preset threshold value. And considering the SNR balancing issue, we set the SNR threshold value of the two transceivers nodes to be the same.

Besides proved that the optimal beamforming matrix of the relay is a 2 by 2 matrix. We also proved that the optimization problem has a set of optimal beamforming solutions and at least one of them is our SOCP solution. The main advantage of SOCP over the existing SDP methods is that it yields a much lower complexity for obtaining the optimal solution.

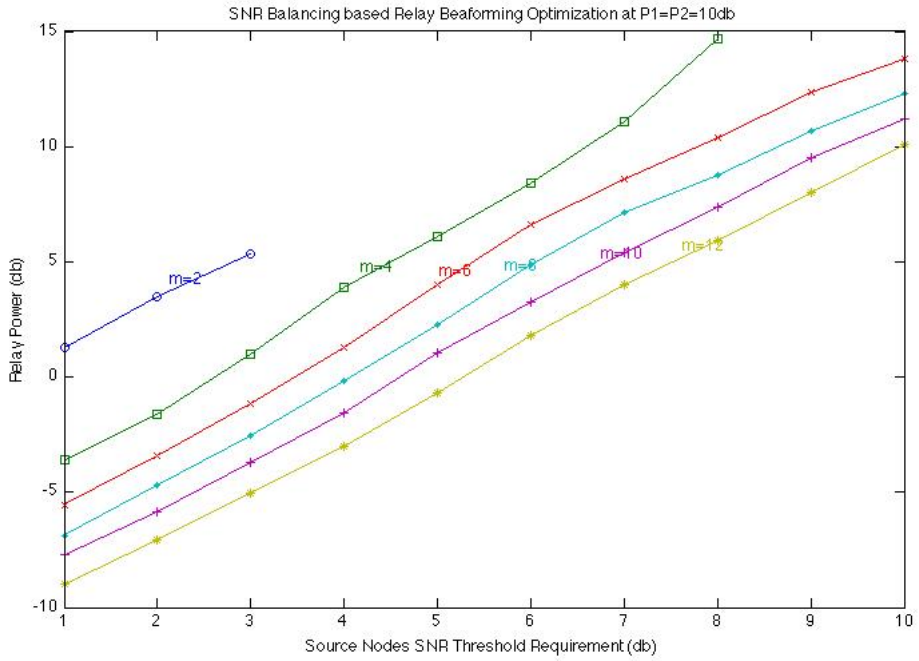

Figure 3.4: Relay Beamforming Optimization at $P_1 = P_2 = 10db$.

Chapter 4

Secrecy Rate Maximisation for TWRC with Perfect CSI

Physical-layer security was first introduced by Wyner *et al.* [87] in 1975 and his results have recently attracted much attention in redefining wireless systems optimization, such as, network secrecy sum-rate maximisation for TWRCs by Mukherjee *et al.* [57], secrecy rate maximisation for OWRCs by Gan *et al.* [95] and joint relay and jammer selection for secure TWRC networks of Chen *et al.* [17]. In this chapter, we revisit the beamforming optimization problem for TWRCs with a multi-antenna AF relay, with the presence of a single antenna eavesdropper. We assume that perfect CSI is available in this network model. We are looking for the optimal solution for minimising the relay power and with the secrecy constraints of the two source nodes. This formulation permits us to keep the messages confidential to the eavesdropper in the information-theoretic sense.

Secrecy related TWRC beamforming optimization is increasingly popular in recent years. A closed-form distributed solution for the TWRC network with multiple single-antenna relays was studied in Wang *et al.* [84]. In this case, the structure of the optimal beamforming matrix can be mathematically formed. Recently, Mo *et al.* [41] discussed the secrecy rate constrained optimization for TWRCs with an untrusted MIMO relay, which is also acted as an eavesdropper. They decomposed this non-convex optimization problem into several sub-problems which can be solved using an iterative algorithm.

Our contribution in this chapter is that the optimal two-way beamforming (TWBF) for minimising the required transmit power of the relay that achieves the target secrecy rates of the two source terminals is determined by solving an SDP in combination with a two-dimensional search. In addition, the optimal two-way zero-forcing (TWZF) solution that places a signal null at the eavesdropper can be obtained from a convex SDP. Results also demonstrate that the optimal TWZF is near-optimal.

4.1 The Wiretap TWRC Model

We consider the TWRC wiretap system as shown in Fig. 4.1 in which we have two transceivers: source nodes S_1 and S_2, each with one antenna, an M-antenna AF-operated relay R, and a single-antenna eavesdropper E. In this network, S_1 and S_2 exchange messages with the aid of R and their messages are required to be kept confidential from E. Communications between S_1 and S_2 takes place in two

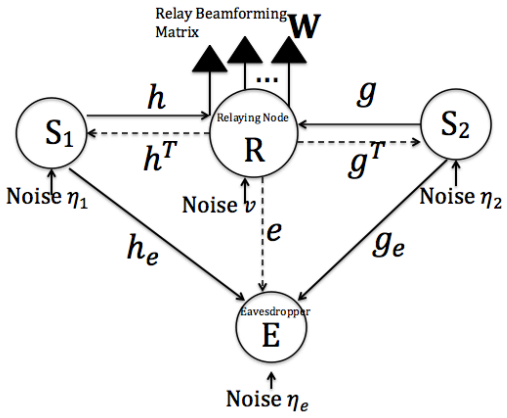

Figure 4.1: The TWRC wiretap model.

consecutive time slots k and $k+1$. This chapters considers just the first two time slots, i.e., $k=1$.

During the *first time slot*, both S_1 and S_2 simultaneously transmit their messages to R. The signals received at R and E can be, respectively, expressed as

$$\mathbf{x} = \sqrt{P_1}\mathbf{h}s_1 + \sqrt{P_2}\mathbf{g}s_2 + \mathbf{v} \text{ (at R)}, \tag{4.1}$$

$$x_e = \sqrt{P_1}h_e s_1 + \sqrt{P_2}g_e s_2 + \eta_e \text{ (at E)}, \tag{4.2}$$

in which P_1 and P_2 are the respective transmit power of S_1 and S_2, s_1 and s_2 are the symbols from S_1 and S_2, respectively, and $\mathbf{v} \in \mathcal{C}^M$ is the additive Gaussian noise vector at R while η_e is the zero-mean Gaussian noise at E such that $\mathbb{E}[\mathbf{v}\mathbf{v}^\dagger] = \sigma^2\mathbf{I}$ and $\mathbb{E}[|\eta_e|^2] = \sigma^2$. The vectors $\mathbf{h} \in \mathcal{C}^M$ and $\mathbf{g} \in \mathcal{C}^M$, are the forward channels from S_1 and S_2 to R, respectively, while the overhearing channels from S_1 and S_2 to E are, respectively, h_e and g_e.

At the *second time slot*, R transmits a beamformed version of $\mathbf{x}$ by a complex weight matrix $\mathbf{W} \in \mathcal{C}^{M\times M}$, $\mathbf{W}\mathbf{x}$, back to S_1 and S_2, which is overheard by E . As such, we have

$$y_1 = \mathbf{h}^T\mathbf{W}\mathbf{x} + \eta_1 \text{ (at } \mathsf{S}_1\text{)} \tag{4.3}$$

$$y_2 = \mathbf{g}^T\mathbf{W}\mathbf{x} + \eta_2 \text{ (at } \mathsf{S}_2\text{)} \tag{4.4}$$

$$y_e = \mathbf{e}^T\mathbf{W}\mathbf{x} + \eta_e' \text{ (at E)} \tag{4.5}$$

in which η_1, η_2 and η_e' denote the respective noises at S_1, S_2 and R and they are assumed to be i.i.d. with zero mean and variance of σ^2. Also, $\mathbf{e} \in \mathcal{C}^M$ denotes the wiretap channel from R to E. In our model, we have assumed that the backward channels from R to S_1 and S_2 are the same as the respective forward channels and they remain static over the period during which our optimization problem is considered.

In TWRCs, s_2 is intended for S_1 with the prior knowledge of its own transmitted message s_1. Therefore, assuming perfect CSI available, S_1 can, after removing the term of s_1, obtain

$$\tilde{y}_1 = \sqrt{P_2}\mathbf{h}^T\mathbf{W}\mathbf{g}s_2 + \mathbf{h}^T\mathbf{W}\mathbf{v} + \eta_1. \tag{4.6}$$

Similarly, for S_2 intending to get s_1, we have

$$\tilde{y}_2 = \sqrt{P_1}\mathbf{g}^T\mathbf{W}\mathbf{h}s_1 + \mathbf{g}^T\mathbf{W}\mathbf{v} + \eta_2. \tag{4.7}$$

At the eavesdropper E, it receives

$$\tilde{e} = \mathbf{e}^T\mathbf{W}\mathbf{X} = \mathbf{e}^T\mathbf{W}(\sqrt{P_1}\mathbf{h}s_1 + \sqrt{P_2}\mathbf{g}s_2 + \mathbf{v} + \eta_e'\sqrt{P_1}\mathbf{e}^T\mathbf{W}\mathbf{h}s_1 + \sqrt{P_2}\mathbf{e}^T\mathbf{W}\mathbf{g}s_2 + \mathbf{e}^T\mathbf{W}\mathbf{v} + \eta_e'.$$

After removal of self interference, it can be seen that the end-to-end channel is reduced to two parallel Gaussian channels. Note that if the self interference is not eliminated, the channel is a broadcast channel (BC). In this case, the achievable coding scheme for the BC (see, e.g., [21] for superposition coding, or [89] for dirty-paper coding) has been well understood.

In the presence of an eavesdropper, a proper performance metric is *secrecy information rate* which measures the achievable rate for error-free communication that is not decodable by the eavesdropper as Wyer *et al.* [87]. The secrecy capacity region for TWRCs is unfortunately unknown. However, as in Mukherjee *et al.* [57, 95], we use the achievable secrecy rate in the form

$$R_{\mathsf{S}} = R_{\mathsf{D}} - R_{\mathsf{E}} \tag{4.8}$$

as the measure of the maximum achievable rate with physical-layer security protection, the difference of the capacity of the main channel and that of the wiretap channel. In particular, R_{E} in (4.8) can be viewed as the capacity of a virtual MIMO multiple-access channel (MAC) which is formed by the wiretapper signal over the two time slots. The same argument has been used in Mukherjee *et al.* [57] *et al.* to justify the metric (4.8). The secrecy rates at S_1 and S_2 are then given, respectively, as

$$\begin{aligned} R_{s1} =& \frac{1}{2}\log_2\left[1 + \frac{P_2|\mathbf{h}^T\mathbf{W}\mathbf{g}|^2}{\sigma^2\left(\|\mathbf{h}^T\mathbf{W}\|^2 + 1\right)}\right] \\ & - \frac{1}{2}\log_2\left[1 + \frac{P_2|\mathbf{e}^T\mathbf{W}\mathbf{g}|^2 + P_2|\mathbf{g}_e|^2}{P_1|\mathbf{e}^T\mathbf{W}\mathbf{h}|^2 + P_1|\mathbf{h}_e|^2 + \sigma^2(\|\mathbf{e}^T\mathbf{W}\|^2 + 2)}\right], \end{aligned} \tag{4.9}$$

$$\begin{aligned} R_{s2} =& \frac{1}{2}\log_2\left[1 + \frac{P_1|\mathbf{g}^T\mathbf{W}\mathbf{h}|^2}{\sigma^2\left(\|\mathbf{g}^T\mathbf{W}\|^2 + 1\right)}\right] \\ & - \frac{1}{2}\log_2\left[1 + \frac{P_1|\mathbf{e}^T\mathbf{W}\mathbf{h}|^2 + P_1|\mathbf{h}_e|^2}{P_2|\mathbf{e}^T\mathbf{W}\mathbf{g}|^2 + P_2|\mathbf{g}_e|^2 + \sigma^2(\|\mathbf{e}^T\mathbf{W}\|^2 + 2)}\right]. \end{aligned} \tag{4.10}$$

4.2 Relay Power Minimisation

Our objective is to minimise the required transmitted power of the relay R subject to individual secrecy rate constraints ε_1 and ε_2, respectively, at S_1 and S_2 by optimising $\mathbf{W}$. That is,

$$\begin{aligned} \min_{\mathbf{W}} \;& P_R \equiv P_1\|\mathbf{W}\mathbf{h}\|^2 + P_2\|\mathbf{W}\mathbf{g}\|^2 + \mathsf{trace}(\mathbf{W}\mathbf{W}^\dagger)\sigma^2 \\ \text{s.t.} \;& \begin{cases} R_{s1} \geq \varepsilon_1, \\ R_{s2} \geq \varepsilon_2. \end{cases} \end{aligned} \tag{4.11}$$

4.2.1 SDP Reformulation

Here, we present an SDP reformulation of (4.11). The rank of optimal matrix is $\mathsf{rank}(\mathbf{W}) = 3$, the detailed proof is provided in the appendix. Now, we define:

$$\mathbf{w} \triangleq \mathsf{vec}(\mathbf{W}) \in \mathcal{C}^{M^2\times 1}, \tag{4.12}$$

$$\mathbf{f}_1 \triangleq \mathsf{vec}(\mathbf{h}\mathbf{g}^T), \tag{4.13}$$

$$\mathbf{f}_2 \triangleq \mathsf{vec}(\mathbf{g}\mathbf{h}^T), \tag{4.14}$$

$$\mathbf{f}_3 \triangleq \mathsf{vec}(\mathbf{e}\mathbf{h}^T), \tag{4.15}$$

$$\mathbf{f}_4 \triangleq \mathsf{vec}(\mathbf{e}\mathbf{g}^T), \tag{4.16}$$

$$\mathbf{U}_1 \triangleq P_1\mathbf{h}\mathbf{h}^\dagger + P_2\mathbf{g}\mathbf{g}^\dagger + \sigma^2\mathbf{I}, \tag{4.17}$$

$$\mathbf{U}_2 \triangleq [\mathsf{diag}(\mathbf{U}_1^T, \mathbf{U}_1^T)]^{\frac{1}{2}}, \tag{4.18}$$

where the notation $\mathsf{vec}(\cdot)$ represents a column vector formed by stacking up all the columns of the input matrix in order. Further, we define the *effective* main and wiretap channels

$$\mathbf{H} \triangleq \boldsymbol{f}(\mathbf{h}) \equiv \mathbf{1}_{M\times 1} \otimes \mathbf{h}, \tag{4.19}$$

$$\mathbf{G} \triangleq \boldsymbol{f}(\mathbf{g}) \equiv \mathbf{1}_{M\times 1} \otimes \mathbf{g}, \tag{4.20}$$

$$\mathbf{E} \triangleq \boldsymbol{f}(\mathbf{e}) \equiv \mathbf{1}_{M\times 1} \otimes \mathbf{e}, \tag{4.21}$$

As such, (4.11) can be rewritten as (4.22),

$$\begin{aligned}
&\min_{\mathbf{W}} \ P_1\|\mathbf{W}\mathbf{h}_1\|^2 + P_2\|\mathbf{W}\mathbf{g}_1\|^2 + \mathsf{trace}(\mathbf{W}\mathbf{B}^\dagger)\sigma^2 \\
&\text{s.t.} \ \begin{cases} \log_2\left[1+\dfrac{P_2|\mathbf{h}^T\mathbf{W}\mathbf{g}_1|^2}{\sigma^2\left(\|\mathbf{h}^T\mathbf{W}\|^2+1\right)}\right] \geq 2\varepsilon_1 + \log_2\left[1+\dfrac{P_2|\mathbf{e}^T\mathbf{W}\mathbf{g}|^2+P_2|\mathbf{g}_e|^2}{P_1|\mathbf{e}^T\mathbf{W}\mathbf{h}_1|^2+P_1|\mathbf{h}_e|^2+\sigma^2(\|\mathbf{e}_1^T\mathbf{W}\|^2+2)}\right], \\ \log_2\left[1+\dfrac{P_1|\mathbf{g}^T\mathbf{W}\mathbf{h}_1|^2}{\sigma^2\left(\|\mathbf{g}^T\mathbf{W}\|^2+1\right)}\right] \geq 2\varepsilon_2 + \log_2\left[1+\dfrac{P_1|\mathbf{e}^T\mathbf{W}\mathbf{h}_1|^2+P_1|\mathbf{h}_e|^2}{P_2|\mathbf{e}^T\mathbf{W}\mathbf{g}_1|^2+P_2|\mathbf{g}_e|^2+\sigma^2(\|\mathbf{e}_1^T\mathbf{W}\|^2+2)}\right]. \end{cases}
\end{aligned} \tag{4.22}$$

Also, by introducing two auxiliary variables

$$t_1 = \frac{P_2|\mathbf{e}^T\mathbf{W}\mathbf{g}|^2+P_2|\mathbf{g}_e|^2}{P_1|\mathbf{e}^T\mathbf{W}\mathbf{h}|^2+P_1|\mathbf{h}_e|^2+\sigma^2(\|\mathbf{e}^T\mathbf{W}\|^2+2)}, \tag{4.23}$$

$$t_2 = \frac{P_1|\mathbf{e}^T\mathbf{W}\mathbf{h}|^2+P_1|\mathbf{h}_e|^2}{P_2|\mathbf{e}^T\mathbf{W}\mathbf{g}|^2+P_2|\mathbf{g}_e|^2+\sigma^2(\|\mathbf{e}^T\mathbf{W}\|^2+2)}, \tag{4.24}$$

and defining

$$\tilde{\varepsilon}_1 \triangleq 2^{2\varepsilon_1+\log_2(1+t_1)} - 1, \tag{4.25}$$

$$\tilde{\varepsilon}_2 \triangleq 2^{2\varepsilon_2+\log_2(1+t_2)} - 1, \tag{4.26}$$

we can rewrite (4.22) as

$$\begin{aligned}
&\min_{\mathbf{w},t_1,t_2} \ \|\mathbf{U}_2\mathbf{w}\|^2 \\
&\text{s.t.} \ \begin{cases}
P_2|\mathbf{f}_1^T\mathbf{w}|^2 \geq \tilde{\varepsilon}_1\sigma^2\left(\|\mathbf{H}\mathbf{w}\|^2+1\right), \\
P_1|\mathbf{f}_2^T\mathbf{w}|^2 \geq \tilde{\varepsilon}_2\sigma^2\left(\|\mathbf{G}\mathbf{w}\|^2+1\right), \\
P_2|\mathbf{f}_4^T\mathbf{w}|^2 = t_1P_1|\mathbf{f}_3^T\mathbf{w}|^2+\sigma^2t_1\left(\|\mathbf{E}\mathbf{w}\|^2+2\right)+t_1P_1|\mathbf{h}_e|^2-P_2|\mathbf{g}_e|^2, \\
P_1|\mathbf{f}_3^T\mathbf{w}|^2 = t_2P_2|\mathbf{f}_4^T\mathbf{w}|^2+\sigma^2t_2\left(\|\mathbf{E}\mathbf{w}\|^2+2\right)+t_2P_2|\mathbf{g}_e|^2-P_1|\mathbf{h}_e|^2.
\end{cases}
\end{aligned} \tag{4.27}$$

We proceed to illustrate that (4.27) has an optimal SDP solution for a given (t_1, t_2). The two equality constraints in (4.27) can be combined into one equality constraint as

$$P_1\left(1+\frac{1}{t_2}\right)|\mathbf{f}_3^T\mathbf{w}|^2-P_2\left(1+\frac{1}{t_1}\right)|\mathbf{f}_4^T\mathbf{w}|^2=P_2\left(1+\frac{1}{t_1}\right)|\mathbf{g}_e|^2-P_1\left(1+\frac{1}{t_2}\right)|\mathbf{h}_e|^2\equiv c. \tag{4.28}$$

Then we define

$$\mathbf{E}_0 \triangleq \mathbf{U}_2^\dagger\mathbf{U}_2, \tag{4.29}$$

$$\mathbf{E}_1 \triangleq \frac{P_2}{\tilde{\varepsilon}_1\sigma^2}\mathbf{f}_1^*\mathbf{f}_1^T-\mathbf{H}^\dagger\mathbf{H}, \tag{4.30}$$

$$\mathbf{E}_2 \triangleq \frac{P_1}{\tilde{\varepsilon}_2\sigma^2}\mathbf{f}_2^*\mathbf{f}_2^T-\mathbf{G}^\dagger\mathbf{G}, \tag{4.31}$$

$$\mathbf{E}_3 \triangleq P_1\left(1+\frac{1}{t_2}\right)\mathbf{f}_3^*\mathbf{f}_3^T-aP_2\left(1+\frac{1}{t_1}\right)\mathbf{f}_4^*\mathbf{f}_4^T. \tag{4.32}$$

Also, we define $\mathbf{X} \triangleq [\mathtt{Re}\{\mathbf{w}\};\mathtt{Im}\{\mathbf{w}\}][\mathtt{Re}\{\mathbf{w}\};\mathtt{Im}\{\mathbf{w}\}]^T$,

$$\mathbf{F}_0 \triangleq \begin{bmatrix} \mathtt{Re}(\mathbf{U}_2) & -\mathtt{Im}(\mathbf{U}_2) \\ \mathtt{Im}(\mathbf{U}_2) & \mathtt{Re}(\mathbf{U}_2) \end{bmatrix}^T \begin{bmatrix} \mathtt{Re}(\mathbf{U}_2) & -\mathtt{Im}(\mathbf{U}_2) \\ \mathtt{Im}(\mathbf{U}_2) & \mathtt{Re}(\mathbf{U}_2) \end{bmatrix}, \tag{4.33}$$

and similarly have $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$, respectively, from $\mathbf{e}$, $\mathbf{E}_2$ and $\mathbf{E}_3$, and $\mathbf{F}_0$ from $\mathbf{U}_2$:

$$\mathbf{F}_1 \triangleq \begin{bmatrix} \mathtt{Re}(\mathbf{E}_1) & -\mathtt{Im}(\mathbf{E}_1) \\ \mathtt{Im}(\mathbf{E}_1) & \mathtt{Re}(\mathbf{E}_1) \end{bmatrix}^T \begin{bmatrix} \mathtt{Re}(\mathbf{E}_1) & -\mathtt{Im}(\mathbf{E}_1) \\ \mathtt{Im}(\mathbf{E}_1) & \mathtt{Re}(\mathbf{E}_1) \end{bmatrix},$$

$$\mathbf{F}_2 \triangleq \begin{bmatrix} \mathtt{Re}(\mathbf{E}_2) & -\mathtt{Im}(\mathbf{E}_2) \\ \mathtt{Im}(\mathbf{E}_2) & \mathtt{Re}(\mathbf{E}_2) \end{bmatrix}^T \begin{bmatrix} \mathtt{Re}(\mathbf{E}_2) & -\mathtt{Im}(\mathbf{E}_2) \\ \mathtt{Im}(\mathbf{E}_2) & \mathtt{Re}(\mathbf{E}_2) \end{bmatrix},$$

$$\mathbf{F}_3 \triangleq \begin{bmatrix} \mathtt{Re}(\mathbf{E}_3) & -\mathtt{Im}(\mathbf{E}_3) \\ \mathtt{Im}(\mathbf{E}_3) & \mathtt{Re}(\mathbf{E}_3) \end{bmatrix}^T \begin{bmatrix} \mathtt{Re}(\mathbf{E}_3) & -\mathtt{Im}(\mathbf{E}_3) \\ \mathtt{Im}(\mathbf{E}_3) & \mathtt{Re}(\mathbf{E}_3) \end{bmatrix}.$$

After rank relaxation, (4.27) becomes

$$\min_{\mathbf{X}} \ \mathsf{trace}(\mathbf{F}_0\mathbf{X}) \quad \text{s.t.} \ \begin{cases} \mathsf{trace}(\mathbf{F}_1\mathbf{X}) \geq 1, \\ \mathsf{trace}(\mathbf{F}_2\mathbf{X}) \geq 1, \\ \mathsf{trace}(\mathbf{F}_3\mathbf{X}) = 0. \end{cases} \tag{4.34}$$

The above is an SDP which can be solved optimally (introduced in dedetails by Boyd *et al.* [14]). As in Zhang *et al.* [93], even if the rank of the optimal solution to (4.34) is greater than one, it will always be possible to obtain a rank-one solution from the higher-rank SDP solution which has the same objective value. Details are given in the appendix.

4.2.2 Optimal TWBF

The optimal TWBF can be obtained by repeatedly solving the SDP in (4.34) in combination with an exhaustive search over (t_1, t_2) according to their definitions. It is easily seen that (4.35) applies,

$$\begin{cases} 0 \leq t_1 \leq \bar{t}_1, \\ 0 \leq t_2 \leq \bar{t}_2, \end{cases} \tag{4.35}$$

where $\bar{t}_1$ and $\bar{t}_2$ are the defined according to:

$$\lambda_{\max}(P_2\mathbf{f}_4^*\mathbf{f}_4^T + P_2|\mathbf{g}_e|^2\mathbf{I}, P_1\mathbf{f}_3^*\mathbf{f}_3^T + \sigma^2\mathbf{E}^\dagger\mathbf{E} + (\sigma^2 + P_1|\mathbf{h}_e|^2)\mathbf{I}) \equiv \bar{t}_1,$$
$$\lambda_{\max}(P_1\mathbf{f}_3^*\mathbf{f}_3^T + P_1|\mathbf{h}_e|^2\mathbf{I}, P_2\mathbf{f}_4^*\mathbf{f}_4^T + \sigma^2\mathbf{E}^\dagger\mathbf{E} + (\sigma^2 + P_2|\mathbf{g}_e|^2)\mathbf{I}) \equiv \bar{t}_2,$$

where $\lambda_{\max}(\mathbf{Z}_1, \mathbf{Z}_2)$ denotes the maximum eigenvalue of the Raleigh quotient, i.e.,

$$\max_{\mathbf{c}} \frac{\mathbf{c}^T\mathbf{Z}_1\mathbf{c}}{\mathbf{c}^T\mathbf{Z}_2\mathbf{c}} \text{ s.t. } \|\mathbf{c}\| \leq 1. \tag{4.36}$$

4.2.3 Optimal TWZF

Due to the complexity of the exhaustive search involved in finding the optimal TWBF, we here propose a low-complexity suboptimal solution based on placing a signal null at E for relay power minimisation. Then the structure of $\mathbf{W}$ is that

$$\mathbf{W} \parallel \mathbf{S} \perp \mathbf{e}$$

and we have $\mathbf{We} = 0$. In this case, the secrecy rates of S_1 and S_2 are, respectively, given by

$$R_{s1} = \frac{1}{2}\log_2\left\{1 + \frac{P_2|\mathbf{h}^T\mathbf{W}\mathbf{g}|^2}{\sigma^2\left(\|\mathbf{h}^T\mathbf{W}\|^2 + 1\right)}\right\}, \tag{4.37}$$

$$R_{s2} = \frac{1}{2}\log_2\left\{1 + \frac{P_1|\mathbf{g}^T\mathbf{W}\mathbf{h}|^2}{\sigma^2\left(\|\mathbf{g}^T\mathbf{W}\|^2 + 1\right)}\right\}. \tag{4.38}$$

which through the null-space aligned zero-forcing method, we have

$$P_2|\mathbf{e}^T\mathbf{W}\mathbf{g}| = P_1|\mathbf{e}^T\mathbf{W}\mathbf{h}| = 0, |\mathbf{f}_3^T\mathbf{w}| \quad = |\mathbf{f}_4^T\mathbf{w}| = 0. \tag{4.39}$$

This means that the beamforming matrix $\mathbf{W}$ is at the null-space of the eavesdropper's channel $\mathbf{e}$. Next, taking the above discussion in to the optimization problem, we have

$$\mathbf{F}\mathbf{w} = 0. \tag{4.40}$$

In this case, (4.27) can be simplified to

$$\min_{\mathbf{w}} \|\mathbf{U}_2\mathbf{w}\|^2 \text{ s.t. } \begin{cases} P_2|\mathbf{f}_1^T\mathbf{w}|^2 \geq \tilde{\varepsilon}_1\sigma^2\left(\|\mathbf{H}\mathbf{w}\|^2+1\right), \\ P_1|\mathbf{f}_2^T\mathbf{w}|^2 \geq \tilde{\varepsilon}_2\sigma^2\left(\|\mathbf{G}\mathbf{w}\|^2+1\right), \\ \quad \mathbf{F}\mathbf{w} = 0. \end{cases} \tag{4.41}$$

Define $\mathbf{F}_4$ from $\mathbf{F}$ as for $\mathbf{F}_0$ from $\mathbf{U}_2$ in (4.33). Then the SDP (4.27) after rank relaxation and the TWZF constraint become

$$\begin{aligned} &\min_{\mathbf{X}} \mathsf{trace}\ (\mathbf{F}_0\mathbf{X}) \\ &\text{s.t.} \begin{cases} \mathsf{trace}(\mathbf{F}_1\mathbf{X}) \geq 1, \\ \mathsf{trace}(\mathbf{F}_2\mathbf{X}) \geq 1, \\ \mathsf{trace}(\mathbf{F}_4\mathbf{X}) = 0. \end{cases} \end{aligned} \tag{4.42}$$

In Appendix I, we prove that the optimal TWZF solution can be exactly found from solving the above SDP and the detailed matrix structure of zero-forcing is discussed in Appendix (A.4).

4.3 Algorithm and Simulation Results

4.3.1 Algorithm

For the TWBF scenario, we use a two-dimensional search algorithm over all possible pairs (t_1, t_2) to achieve the minimum P_T. Define

$$t_{1,u} = \lambda_{\max}(P_2\mathbf{f}_4^*\mathbf{f}_4^T + P_2|\mathbf{g}_e|^2\mathbf{I}, P_1\mathbf{f}_3^*\mathbf{f}_3^T + \sigma^2\mathbf{E}^\dagger\mathbf{E} + (\sigma^2 + P_1|\mathbf{h}_e|^2)\mathbf{I}), \tag{4.43}$$

$$t_{2,u} = \lambda_{\max}(P_1\mathbf{f}_3^*\mathbf{f}_3^T + P_1|\mathbf{h}_e|^2\mathbf{I}, P_2\mathbf{f}_4^*\mathbf{f}_4^T + \sigma^2\mathbf{E}^\dagger\mathbf{E} + (\sigma^2 + P_2|\mathbf{g}_e|^2)\mathbf{I}). \tag{4.44}$$

Also, define the resolution $\Delta t = \frac{t_{1,u}}{N}$ for some large N as the iteration index. For $\mathbf{X}$ corresponding to a given $t_{1,u}$, we can compute a corresponding t_2, denoted as $t_{2,l}$. Taking $t_1 = t_{1,u}, t_2 = t_{2,l}$ into (A-33), we get an optimal objective value and can serve it as our achievable relay power $P_{T,\mathsf{opt}}$. Initialise $t_{1,o} = t_{1,u}, t_{2,o} = t_{2,l}$, and $i = N$ as the iteration index and $\Delta t = \frac{t_{1,u}}{N}$. We then use a brute force strategy to check each point iteratively staring from $t_{1,u}$ down to 0. In each iteration, (A-33) is an SDP, and we can use the bisection search over t_2 to check (A-33) and at the end the bisection search will provide a smaller objective value than the one saved before. Thus, the value at the end of the algorithm

is a global minimum of the relay power based on the secrecy rate constraints.

Algorithm 1 Proposed 2D Bisection Algorithm

Initialize $t_{1,min} = t_{2,min} = 0$,$t_{1,max} = t_{1,u}$, $t_{2,max} = t_{2,u}$;
while $t_{1,max} - t_{1,min} \geq \varepsilon$ **do**
  while $t_{2,max} - t_{2,min} \geq \varepsilon$ **do**
    $t_2 = \frac{t_{2,min}+t_{2,max}}{2}$.
    Solve problem (A-33) with the above t_1 and t_2 and get optimal objective value P_T
    Solve problem (A-33) with the above t_1 and $t_2 = t_2 + \Delta t$ for $\Delta t > 0$ and get optimal objective value $P_T^{'}$
    if $P_T > P_T^{'}$ **then**
      $t_{2,min} = t_2$
    else
      $t_{2,max} = t_2$
    end if
  end while
  Set $t_1 = t_1 + \Delta t$ with $\Delta t > 0$
  while $t_{2,max} - t_{2,min} \geq \varepsilon$ **do**
    $t_2 = \frac{t_{2,min}+t_{2,max}}{2}$.
    Solve problem (A-33) with the above t_1 and t_2 and get optimal objective value P_{T1}
    Solve problem (A-33) with the above t_1 and $t_2 = t_2 + \Delta t$ for $\Delta t > 0$ and get optimal objective value P_{T1}
    if $P_{T1} > P_{T1}^{'}$ **then**
      $t_{2,min} = t_2$
    else
      $t_{2,max} = t_2$
    end if
  end while
  if $P_{T1} > P_{T1}^{'}$ **then**
    $t_{1,min} = t_1$
  else
    $t_{1,max} = t_1$
  end if
end while

4.3.2 Analysis and Simulation

Simulations are conducted to evaluate the performance of the proposed systems. In the simulations, it is assumed that $\frac{P_1}{\sigma^2} = \frac{P_2}{\sigma^2} = 10$ and that the channels are all complex Gaussian with zero means and unit-variances. Results in Fig. 4.2 are provided for the average transmit relay SNR against the secrecy rate requirements, $\varepsilon_1 = \varepsilon_2 = \varepsilon > 0$, for various number of antennas at the relay, M. Results indicate that the performance of TWZF is very close to that of the optimal TWBF and the required SNR increases with the target secrecy rate. On the other hand, results in Fig. 4.3 show the probability of the problem being feasible or not. As expected, we see that the probability that the given target secrecy rates are feasible decreases if the targets increase.

4.4 Conclusion

In this chapter, beamforming for AF relaying has been studied under secrecy constraints in TWRC. The relay of this TWRC is a MIMO really which performs beamforming function through beamforming matrix $\mathbf{W}$. We consider that the single secrecy rate for each transceiver nodes S_1 and S_2, with the

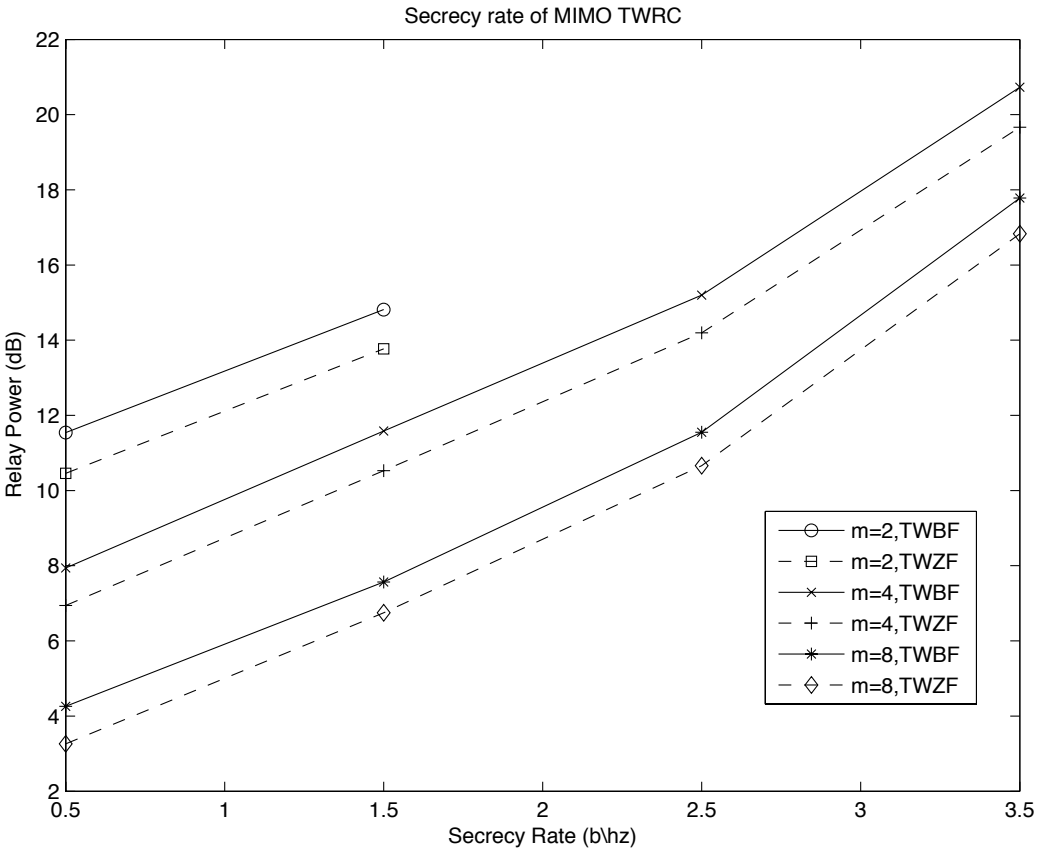

Figure 4.2: Transmit relay SNR versus the secrecy rate requirements.

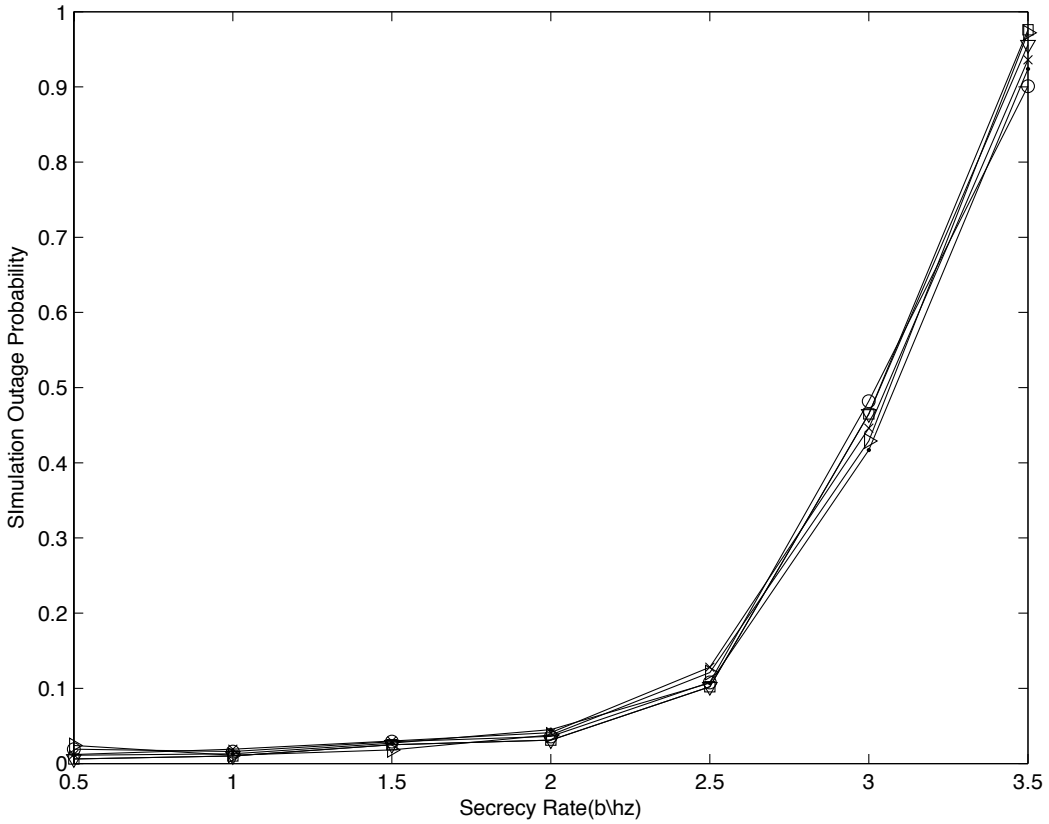

Figure 4.3: Outage probability of secrecy rate.

existence of an extra eavesdropper E. A optimisation problem is formed with the objective to minimize the relay power and two constraints that single secrecy rates of S_1 and S_2 are satisfied with a pre-set threshold positive value. Besides, we also considered that the power allocated to S_1 and S_2 and the same. We assumed that all the nodes within the network have the perfect CSI of the others as well.

For the original non-convex optimization problem, we have presented an SDP formulation based 2D searching algorithm, under which optimal beamforming designs that maximise secrecy rates were provided under the relay power constraints, if perfect CSI is exploited. A low-complex TWZF solution based on generalised Raleigh quotient was also obtained achieving near-optimal performance.

Chapter 5

Robust Secure TWRC Beamforming

In this chapter, we revisit the same optimization problem of relay beamforming for TWRC in Chapter 5 except that we have this time ICSI where the errors are assumed bounded by ellipsoids. Our contribution is the maxmisation of the sum secrecy rate under the total power constraints by solving an SDP after transferring the optimization constraints into LMIs, under a rank relaxation condition.

5.1 Network Model

Consider the same network model as in Chapter 5, S_1 and S_2 exchange messages with the aid of R and their messages are required to be kept confidential from E. Communications between S_1 and S_2 takes place in every two consecutive time slots. As before, after two time slots of transmission, we have

$$y_1 = \mathbf{h}^T\mathbf{W}\mathbf{x} + \eta_1 \ (\text{at } \mathsf{S}_1), \tag{5.1}$$

$$y_2 = \mathbf{g}^T\mathbf{W}\mathbf{x} + \eta_2 \ (\text{at } \mathsf{S}_2), \tag{5.2}$$

$$y_e = \mathbf{e}^T\mathbf{W}\mathbf{x} + \eta_e \ (\text{at } \mathsf{E}), \tag{5.3}$$

where η_1, η_2 and η_e denote the respective noise at S_1, S_2 and R and they are assumed to be i.i.d. with zero mean and variance of σ^2. Also, $\mathbf{e} \in \mathcal{C}^M$ denotes the wiretap channel from R to E. In our model, same as before, we have assumed that the backward channels from R to S_1 and S_2 are the same as the respective forward channels and they remain static over the period of optimization of interest.

However, the CSI is considered to be imperfect and modelled as

$$\mathbf{h} = \hat{\mathbf{h}} + \tilde{\mathbf{h}}, \tag{5.4}$$

$$\mathbf{g} = \hat{\mathbf{g}} + \tilde{\mathbf{g}}, \tag{5.5}$$

$$\mathbf{e} = \hat{\mathbf{e}} + \tilde{\mathbf{e}}, \tag{5.6}$$

where $\mathbf{h}$ and $\mathbf{g}$ are the true CSI, $\hat{\mathbf{h}}$, $\hat{\mathbf{g}}$ and $\hat{\mathbf{e}}$ are the imperfect CSI available at the relay nodes, and $\tilde{\mathbf{h}}$, $\tilde{\mathbf{g}}$ and $\tilde{\mathbf{e}}$, represent the additive errors in the CSI. Further, we assume that $\|\tilde{\mathbf{h}}\| \leq \varepsilon_{\mathbf{h}}$, $\|\tilde{\mathbf{g}}\| \leq \varepsilon_{\mathbf{g}}$ and

$\|\tilde{\mathbf{e}}\| \leq \varepsilon_{\mathbf{e}}$. In other words, $\mathbf{h}$, $\mathbf{g}$ and $\mathbf{e}$ belong to the uncertainty sets $\mathscr{R}_h$, $\mathscr{R}_g$ and $\mathscr{R}_e$, respectively,

$$\mathscr{R}_h = \left\{\zeta | \zeta = \hat{\mathbf{h}} + \tilde{\mathbf{h}}, \|\tilde{\mathbf{h}}\| \leq \varepsilon_{\mathbf{h}}\right\}, \tag{5.7}$$

$$\mathscr{R}_g = \left\{\zeta | \zeta = \hat{\mathbf{g}} + \tilde{\mathbf{g}}, \|\tilde{\mathbf{g}}\| \leq \varepsilon_{\mathbf{g}}\right\}, \tag{5.8}$$

$$\mathscr{R}_e = \left\{\zeta | \zeta = \hat{\mathbf{e}} + \tilde{\mathbf{e}}, \|\tilde{\mathbf{e}}\| \leq \varepsilon_{\mathbf{e}}\right\}. \tag{5.9}$$

Taking this imperfect CSI model into (5.1)–(4.5), we have

$$\begin{aligned} y_1 =& \sqrt{P_1}(\hat{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{h}} + \tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{h}} + \tilde{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{h}})s_1 \\ &+ \sqrt{P_2}(\hat{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}})s_2 + (\hat{\mathbf{h}}^T\mathbf{W} + \tilde{\mathbf{h}}^T\mathbf{W})\mathbf{v} + \eta_1, \end{aligned} \tag{5.10}$$

$$\begin{aligned} y_2 =& \sqrt{P_1}(\hat{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}})s_1 \\ &+ \sqrt{P_2}(\hat{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{g}} + \tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{g}} + \tilde{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{g}})s_2 + (\hat{\mathbf{g}}^T\mathbf{W} + \tilde{\mathbf{g}}^T\mathbf{W})\mathbf{v} + \eta_2, \end{aligned} \tag{5.11}$$

$$y_e = \sqrt{P_1}\mathbf{e}^T\mathbf{W}(\hat{\mathbf{h}} + \tilde{\mathbf{h}})s_1 + \sqrt{P_2}\mathbf{e}^T\mathbf{W}(\hat{\mathbf{g}} + \tilde{\mathbf{g}})s_1 + \mathbf{e}^T\mathbf{W}\mathbf{v} + \eta_e. \tag{5.12}$$

5.2 Problem Formulation

In TWRCs, s_2 is intended for S_1 with the prior knowledge of its own transmitted message s_1. The influence of the second order CSI errors compared to the first-order is negligible. Therefore, with the imperfect CSI model, we can obtain

$$\begin{aligned} \tilde{y}_1 =& \underbrace{\sqrt{P_1}(\tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{h}})s_1}_{\text{remaining-self-interference}} \\ &+ \underbrace{\sqrt{P_2}(\hat{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}})s_2}_{\text{designed-signal}} + \underbrace{(\hat{\mathbf{h}}\mathbf{W} + \tilde{\mathbf{h}}\mathbf{W})\mathbf{v} + \eta_1}_{\text{noise}}. \end{aligned} \tag{5.13}$$

Similarly, for S_2 intending to get s_1, we have

$$\begin{aligned} \tilde{y}_2 =& \underbrace{\sqrt{P_2}(\tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{g}})s_2}_{\text{remaining-self-interference}} \\ &+ \underbrace{\sqrt{P_1}(\hat{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}})s_1}_{\text{designed-signal}} + \underbrace{(\hat{\mathbf{g}}^T\mathbf{W} + \tilde{\mathbf{g}}^T\mathbf{W})\mathbf{v} + \eta_2}_{\text{noise}}. \end{aligned} \tag{5.14}$$

At the eavesdropper E, it receives

$$y_e = \underbrace{\sqrt{P_1}(\hat{\mathbf{e}} + \tilde{\mathbf{e}})^T\mathbf{W}(\hat{\mathbf{h}} + \tilde{\mathbf{h}})s_1 + \sqrt{P_2}(\hat{\mathbf{e}} + \tilde{\mathbf{e}})^T\mathbf{W}(\hat{\mathbf{g}} + \tilde{\mathbf{g}})s_1}_{\text{signal}} + \underbrace{(\hat{\mathbf{e}} + \tilde{\mathbf{e}})^T\mathbf{W}\mathbf{v} + \eta_e}_{\text{noise}}. \tag{5.15}$$

Once the self interference is removed at both terminals, we can see that the end-to-end channel now is essentially two parallel Gaussian channels. If the self interference were still present, the channel is still the BC. This is because in the achievable coding scheme for the BC (see, for example [21] for

superposition coding, or [89] for DPC), it is assumed that the broadcast signal is made of several parts, and each part is meant for a terminal to decode. This enables us to view the broadcast signal from the viewpoint, as being made up of the part destined for that terminal to decode plus interference.

As in [57, 95], we use the achievable secrecy rate in the form

$$R_{\mathsf{S}} = R_{\mathsf{D}} - R_{\mathsf{E}} \tag{5.16}$$

as the measure of the maximum achievable rate with physical-layer security protection, which is the difference of the capacity of the main channel and that of the wiretap channel. R_{E} is formed by forming the wiretapper signal over 2 time slots and it considers the effect of the wiretapper over 2 time slots as a virtual MAC over the 2 time slots. R_{E} is then found by the capacity of the MIMO-MAC. Details using this argument can be found in the work by Mukherjee *et al.* in [57]. The perfect secrecy rates at S_1 and S_2 are given, respectively, as

$$R_1 = \frac{1}{2}\log\left(1+\mathsf{SNR}_1\right) - \frac{1}{2}\log\left(1+\mathsf{SNR}_{e1}\right), \tag{5.17}$$

$$R_2 = \frac{1}{2}\log\left(1+\mathsf{SNR}_2\right) - \frac{1}{2}\log\left(1+\mathsf{SNR}_{e2}\right). \tag{5.18}$$

However, the secrecy capacity region for each end node in the TWRCs is still unknown. Obviously, it is relevant to consider the case that the eavesdropper is trying to decode both the signals s_1 and s_2. Therefore, the information sum-rate achieved at the eavesdropper over the 2 time slots can be expressed as the maximum sum-rate of a two-user MIMO MAC.

Here, we consider the case that the eavesdropper performs joint decoding from

$$\begin{bmatrix} \mathbf{x}_e \\ \mathbf{y}_e \end{bmatrix} = \begin{bmatrix} h_e & g_e \\ \sqrt{P_1}\mathbf{e}^T\mathbf{W}\mathbf{h} & \sqrt{P_2}\mathbf{e}^T\mathbf{W}\mathbf{g} \end{bmatrix} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} + \begin{bmatrix} \eta_e \\ \mathbf{e}^T\mathbf{W} + \eta_e' \end{bmatrix}. \tag{5.19}$$

In [79], the achievable secrecy rate region for using Gaussian inputs and stochastic encoders for the wiretap TWRC is given as

$$\sum_n R_n^s \leq \left[\sum_n I(y_n; s_n) - I(y_e; s_n)\right]. \tag{5.20}$$

Given that the eavesdropper performs joint decoding, the achievable sum secrecy rate can be found as

$$R_{\mathsf{sum}}^s = \left[I(\tilde{y}_1; s_2) + I(\tilde{y}_2; s_1) - I(\tilde{y}_e; s_1, s_2)\right]^+, \tag{5.21}$$

where $[a]^+ = \max(0, a)$.

The eavesdropper's rate based on joint decoding is given by

$$R_e = \frac{1}{2}\log\left(1 + \frac{\begin{bmatrix} h_e & g_e \\ \sqrt{P_1}\mathbf{e}^T\mathbf{W}\mathbf{h} & \sqrt{P_2}\mathbf{e}^T\mathbf{W}\mathbf{g} \end{bmatrix}^\dagger \begin{bmatrix} h_e & g_e \\ \sqrt{P_1}\mathbf{e}^T\mathbf{W}\mathbf{h} & \sqrt{P_2}\mathbf{e}^T\mathbf{W}\mathbf{g} \end{bmatrix}}{\begin{bmatrix} \eta_e \\ \mathbf{e}^T\mathbf{W} + \eta_e' \end{bmatrix}^\dagger \begin{bmatrix} \eta_e \\ \mathbf{e}^T\mathbf{W} + \eta_e' \end{bmatrix}}\right) \tag{5.22}$$

and

$$R_s = R_1 + R_2 - R_e = \frac{1}{2}\log\left(1 + \mathsf{SNR}_1\right) + \frac{1}{2}\log\left(1 + \mathsf{SNR}_2\right) - \frac{1}{2}\log\left(1 + \mathsf{SNR}_e\right), \tag{5.23}$$

where

$$\mathsf{SNR}_1 = \frac{P_2|\hat{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}}|^2}{P_1|\tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{h}}|^2 + \sigma^2(\|\hat{\mathbf{h}}^T\mathbf{W} + \tilde{\mathbf{h}}^T\mathbf{W}\|^2 + 1)}, \tag{5.24}$$

$$\mathsf{SNR}_2 = \frac{P_1|\hat{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}}|^2}{P_2|\tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{g}}|^2 + \sigma^2(\|\hat{\mathbf{g}}^T\mathbf{W} + \tilde{\mathbf{g}}^T\mathbf{W}\|^2 + 1)}, \tag{5.25}$$

$$\mathsf{SNR}_e = \frac{P_1|(\hat{\mathbf{e}} + \tilde{\mathbf{e}})^T\mathbf{W}(\hat{\mathbf{h}} + \tilde{\mathbf{h}})|^2 + P_2|(\hat{\mathbf{e}} + \tilde{\mathbf{e}})^T\mathbf{W}(\hat{\mathbf{g}} + \tilde{\mathbf{g}})|^2 + P_1|h_e|^2 + P_2|g_e|^2}{\sigma^2(\|\hat{\mathbf{e}} + \tilde{\mathbf{e}}^T\mathbf{W}\|^2 + 2)}. \tag{5.26}$$

Our objective is to maximise the sum secrecy rate R_s subject to the total power constraint, i.e.,

$$\max_{\mathbf{W}} R_s \;\text{ s.t. }\; P_1\|\mathbf{W}\mathbf{h}\|^2 + P_2\|\mathbf{W}\mathbf{g}\|^2 + \mathsf{trace}(\mathbf{W}\mathbf{W}^\dagger)\sigma^2 \le P_R. \tag{5.27}$$

5.3 An SDP Solution

Our objective is to transfer the constraint of the above optimization problem into LMIs. To do so, we find it useful to first define the following list of vectors:

$$\bar{\mathbf{h}} = \hat{\mathbf{h}} \otimes \mathbf{1}_{M\times 1} = \begin{bmatrix} \hat{h}_1\mathbf{1}_M \\ \vdots \\ \hat{h}_m\mathbf{1}_M \end{bmatrix}, \tag{5.28}$$

$$\check{\mathbf{h}} = \mathbf{1}_{M\times 1} \otimes \hat{\mathbf{h}} = \begin{bmatrix} \hat{\mathbf{h}} \\ \vdots \\ \hat{\mathbf{h}} \end{bmatrix}; \tag{5.29}$$

$$\Delta\bar{\mathbf{h}} = \tilde{\mathbf{h}} \otimes \mathbf{1}_{M\times 1} = \begin{bmatrix} \tilde{h}_1\mathbf{1}_M \\ \vdots \\ \tilde{h}_m\mathbf{1}_M \end{bmatrix}, \tag{5.30}$$

$$\Delta\check{\mathbf{h}} = \mathbf{1}_{M\times 1} \otimes \tilde{\mathbf{h}} = \begin{bmatrix} \tilde{\mathbf{h}} \\ \vdots \\ \tilde{\mathbf{h}} \end{bmatrix}, \tag{5.31}$$

$$\bar{\mathbf{g}} = \hat{\mathbf{g}} \otimes \mathbf{1}_{M\times 1} = \begin{bmatrix} \hat{g}_1\mathbf{1}_M \\ \vdots \\ \hat{g}_m\mathbf{1}_M \end{bmatrix}, \tag{5.32}$$

$$\check{\mathbf{g}} = \mathbf{1}_{M\times 1} \otimes \hat{\mathbf{g}} = \begin{bmatrix} \hat{\mathbf{g}} \\ \vdots \\ \hat{\mathbf{g}} \end{bmatrix}, \tag{5.33}$$

$$\Delta\bar{\mathbf{g}} = \tilde{\mathbf{g}} \otimes \mathbf{1}_{M\times 1} = \begin{bmatrix} \tilde{g}_1\mathbf{1}_M \\ \dots \\ \tilde{g}_m\mathbf{1}_M \end{bmatrix}, \tag{5.34}$$

$$\Delta\check{\mathbf{g}} = \mathbf{1}_{M\times 1} \otimes \tilde{\mathbf{g}} = \begin{bmatrix} \tilde{\mathbf{g}} \\ \vdots \\ \tilde{\mathbf{g}} \end{bmatrix}, \tag{5.35}$$

$$\bar{\mathbf{e}} = \hat{\mathbf{e}} \otimes \mathbf{1}_{M\times 1} = \begin{bmatrix} \hat{e}_1\mathbf{1}_M \\ \vdots \\ \hat{e}_m\mathbf{1}_M \end{bmatrix}, \tag{5.36}$$

$$\check{\mathbf{e}} = \mathbf{1}_{M\times 1} \otimes \hat{\mathbf{e}} = \begin{bmatrix} \hat{\mathbf{e}} \\ \vdots \\ \hat{\mathbf{e}} \end{bmatrix}, \tag{5.37}$$

$$\Delta\bar{\mathbf{e}} = \tilde{\mathbf{e}} \otimes \mathbf{1}_{M\times 1} = \begin{bmatrix} \tilde{e}_1\mathbf{1}_M \\ \vdots \\ \tilde{e}_m\mathbf{1}_M \end{bmatrix}, \tag{5.38}$$

$$\Delta\check{\mathbf{e}} = \mathbf{1}_{M\times 1} \otimes \tilde{\mathbf{e}} = \begin{bmatrix} \tilde{\mathbf{e}} \\ \vdots \\ \tilde{\mathbf{e}} \end{bmatrix}. \tag{5.39}$$

Instead of treating the CSI errors of $\mathbf{h}$,$\mathbf{h}$ and $\mathbf{e}$ separately, we combine them by defining:

$$\mathbf{c} = \mathsf{vec}(\mathbf{h}, \mathbf{g}, \mathbf{e}), \tag{5.40}$$

$$\Delta\mathbf{c} = \mathsf{vec}(\tilde{\mathbf{h}}, \tilde{\mathbf{g}}, \tilde{\mathbf{e}}). \tag{5.41}$$

In addition, to reach the SDP form of the optimization problem, we also define the following parameters:

$$\mathbf{w} = \mathsf{vec}(\mathbf{W}), \tag{5.42}$$

$$\bar{\mathbf{W}} = \mathbf{w}\mathbf{w}^{\dagger}, \tag{5.43}$$

$$\mathbf{G}_1 = \begin{bmatrix} \mathbf{I}_M & \mathbf{O}_M & \mathbf{O}_M \end{bmatrix} \in \mathcal{R}^{M\times 3M}, \tag{5.44}$$

$$\mathbf{G}_2 = \begin{bmatrix} \mathbf{O}_M & \mathbf{I}_M & \mathbf{O}_M \end{bmatrix} \in \mathcal{R}^{M\times 3M}, \tag{5.45}$$

$$\mathbf{G}_3 = \begin{bmatrix} \mathbf{O}_M & \mathbf{O}_M & \mathbf{I}_M \end{bmatrix} \in \mathcal{R}^{M\times 3M}, \tag{5.46}$$

$$\mathbf{D}_R = \mathbf{I}_M \otimes \mathbf{1}_{M\times 1} = \begin{bmatrix} \mathbf{1}_{M\times 1} & \mathbf{0}_{M\times 1} & \dots & \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} & \mathbf{1}_{M\times 1} & \dots & \mathbf{0}_{M\times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{M\times 1} & \mathbf{0}_{M\times 1} & 0 & \mathbf{1}_{M\times 1} \end{bmatrix} \in \mathcal{R}^{M^2\times M}, \tag{5.47}$$

$$\mathbf{D}_l = \mathbf{1}_{M\times 1} \otimes \mathbf{I}_M = \begin{bmatrix} \mathbf{I}_M \\ \mathbf{I}_M \\ \vdots \\ \mathbf{I}_M \end{bmatrix} \in \mathcal{R}^{M^2\times M}, \tag{5.48}$$

where $\mathbf{O}_M$ is an all-zero $M \times M$ matrix , and $\mathbf{I}_M$ denotes an $M \times M$ identity matrix. Using the above definitions, we can see that

$$\begin{aligned} \mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c} &= \begin{bmatrix} \mathbf{1}_{M\times 1} & \mathbf{0}_{M\times 1} & \dots & \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} & \mathbf{1}_{M\times 1} & \dots & \mathbf{0}_{M\times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{M\times 1} & \mathbf{0}_{M\times 1} & 0 & \mathbf{1}_{M\times 1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_M & \mathbf{O}_M & \mathbf{O}_M \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{h}} \\ \tilde{\mathbf{g}} \\ \tilde{\mathbf{e}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{1}_{M\times 1} & \mathbf{0}_{M\times 1} & \dots & \mathbf{0}_{M\times 1} \\ \mathbf{0}_{M\times 1} & \mathbf{1}_{M\times 1} & \dots & \mathbf{0}_{M\times 1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{M\times 1} & \mathbf{0}_{M\times 1} & 0 & \mathbf{1}_{M\times 1} \end{bmatrix} \tilde{\mathbf{h}} = \begin{bmatrix} \tilde{h}_1 \\ \vdots \\ \tilde{h}_1 \\ \vdots \\ \tilde{h}_m \\ \vdots \\ \tilde{h}_m \end{bmatrix} = \tilde{\mathbf{h}} \otimes \mathbf{1}_{M\times 1}, \end{aligned} \tag{5.49}$$

and

$$\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c} = \begin{bmatrix} \mathbf{I}_M \\ \mathbf{I}_M \\ \vdots \\ \mathbf{I}_M \end{bmatrix} \begin{bmatrix} \mathbf{I}_M & \mathbf{O}_M & \mathbf{O}_M \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{h}} \\ \tilde{\mathbf{g}} \\ \tilde{\mathbf{e}} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_M \\ \mathbf{I}_M \\ \vdots \\ \mathbf{I}_M \end{bmatrix} \tilde{\mathbf{h}} = \begin{bmatrix} \tilde{\mathbf{h}} \\ \vdots \\ \tilde{\mathbf{h}} \end{bmatrix} = \mathbf{1}_{M\times 1} \otimes \tilde{\mathbf{h}}. \tag{5.50}$$

Similar to (5.49) and (5.50) above, we can also rewrite the channel uncertainty parameters as

$$\Delta\bar{\mathbf{h}} = \tilde{\mathbf{h}} \otimes \mathbf{1}_{M\times 1} = \mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}; \tag{5.51}$$

$$\Delta\check{\mathbf{h}} = \mathbf{1}_{M\times 1} \otimes \tilde{\mathbf{h}} = \mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c}; \tag{5.52}$$

$$\Delta\bar{\mathbf{g}} = \tilde{\mathbf{g}} \otimes \mathbf{1}_{M\times 1} = \mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}; \tag{5.53}$$

$$\Delta\check{\mathbf{g}} = \mathbf{1}_{M\times 1} \otimes \tilde{\mathbf{g}} = \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c}; \tag{5.54}$$

$$\Delta\bar{\mathbf{e}} = \tilde{\mathbf{e}} \otimes \mathbf{1}_{M\times 1} = \mathbf{D}_R\mathbf{G}_3\Delta\mathbf{c}; \tag{5.55}$$

$$\Delta\check{\mathbf{e}} = \mathbf{1}_{M\times 1} \otimes \tilde{\mathbf{e}} = \mathbf{D}_L\mathbf{G}_3\Delta\mathbf{c}. \tag{5.56}$$

Now, we consider the signal part of SNR_1, which can be re-expressed as

$$\mathbf{h}^T\mathbf{W}\mathbf{g} = \sum_{i,j=1}^{m} h_i g_j w_{ij} = \begin{bmatrix} \tilde{g}_1 \\ \vdots \\ \tilde{g}_1 \\ \vdots \\ \tilde{g}_m \\ \vdots \\ \tilde{g}_m \end{bmatrix} \otimes \begin{bmatrix} \tilde{h}_1 \\ \vdots \\ \tilde{h}_m \\ \vdots \\ \tilde{h}_1 \\ \vdots \\ \tilde{h}_m \end{bmatrix} \mathsf{vec}(\mathbf{W}) = [(\mathbf{g} \otimes \mathbf{1}_{M\times 1}) \odot (\mathbf{1}_{M\times 1}) \otimes \mathbf{h})]^T \mathsf{vec}(\mathbf{W}). \tag{5.57}$$

Also, we can rewire the numerator of SNR_1 and denote it as a_1:

$$\begin{aligned} a_1 =& P_2|\hat{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}}|^2 \\ =& P_2|(\bar{\mathbf{g}} \odot \check{\mathbf{h}} + \Delta\bar{\mathbf{g}} \odot \check{\mathbf{h}} + \bar{\mathbf{g}} \odot \Delta\check{\mathbf{h}})^T\mathbf{w}|^2 \\ =& P_2|(\bar{\mathbf{g}} \odot \check{\mathbf{h}} + \mathsf{diag}(\check{\mathbf{h}})\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c} + \mathsf{diag}(\bar{\mathbf{g}})\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c})^T\mathbf{w}|^2 \\ =& P2\Bigg((\bar{\mathbf{g}} \odot \check{\mathbf{h}})^T\bar{\mathbf{W}}(\bar{\mathbf{g}} \odot \check{\mathbf{h}})^* \\ &+ \Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c} + \Delta\mathbf{c}^T\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c} \\ &+ \Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c} + \Delta\mathbf{c}^T\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c} \\ &+ 2(\bar{\mathbf{g}} \odot \check{\mathbf{h}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}^* + 2(\bar{\mathbf{g}} \odot \check{\mathbf{h}})^T\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c}^* \Bigg). \end{aligned} \tag{5.58}$$

Then we further define

$$\begin{aligned} \mathbf{Q}_1 =& P_2\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1 + P_2\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_2 \\ &+ P_2\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_2 + P_2\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1; \end{aligned} \tag{5.59}$$

$$\mathbf{q}_1^\dagger = P_2(\bar{\mathbf{g}} \odot \check{\mathbf{h}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_2 + P_2(\bar{\mathbf{g}} \odot \check{\mathbf{h}})^T\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1; \tag{5.60}$$

$$k_1 = P_2(\bar{\mathbf{g}} \odot \check{\mathbf{h}})^T\bar{\mathbf{W}}(\bar{\mathbf{g}} \odot \check{\mathbf{h}})^*. \tag{5.61}$$

As a consequence, we have

$$a_1 = \Delta\mathbf{c}^T\mathbf{Q}_1\Delta\mathbf{c}^* + 2\mathtt{Re}(\mathbf{q}_1^\dagger\Delta\mathbf{c}^*) + k_1. \tag{5.62}$$

For the denominator of SNR_1, denoted as b_1, we have

$$\begin{aligned} b_1 &= P_1|\tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{h}}|^2 + \sigma^2(\|\hat{\mathbf{h}}^T\mathbf{W} + \tilde{\mathbf{h}}^T\mathbf{W}\|^2 + 1) \\ &= P_1|(\bar{\mathbf{h}}\odot\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c})^T\mathbf{w} + (\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}\odot\check{\mathbf{h}})^T\mathbf{w}|^2 \\ &\quad + \sigma^2\left(\|\hat{\mathbf{h}}^T\mathbf{W}\|^2 + \|\tilde{\mathbf{h}}^T\mathbf{W}\|^2 + 2\mathtt{Re}(\hat{\mathbf{h}}^T\mathbf{W}\mathbf{W}^\dagger\tilde{\mathbf{h}}^*) + 1\right). \end{aligned} \tag{5.63}$$

Similary, we define

$$\begin{aligned} \mathbf{Q}_2 =& P_1\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_1 + P_1\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_1 \\ &+ P_1\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_1 + P_1\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_1 \\ &+ \sigma^2\mathbf{G}_1^T\mathbf{D}_L^T\mathbf{E}\odot\bar{\mathbf{W}}\mathbf{D}_L\mathbf{G}_1; \end{aligned} \tag{5.64}$$

$$\mathbf{q}_2^\dagger = \sigma^2\Delta\check{\mathbf{h}}^T\mathbf{E}\odot\bar{\mathbf{W}}\mathbf{D}_L\mathbf{G}_1; \tag{5.65}$$

$$k_2 = \sigma^2\left(\check{\mathbf{h}}^T\mathbf{E}\odot\bar{\mathbf{W}}\check{\mathbf{h}}^* + 1\right), \tag{5.66}$$

where $\mathbf{E} = \mathbf{I}_M \otimes (\mathbf{1}_{M\times 1}\mathbf{1}_{M\times 1}^T)$. Consequently, we get

$$b_1 = \Delta\mathbf{c}^T\mathbf{Q}_2\Delta\mathbf{c}^* + 2\mathtt{Re}(\mathbf{q}_2^\dagger\Delta\mathbf{c}^*) + k_2. \tag{5.67}$$

Furthermore, we can also define:

$$\begin{aligned} \mathbf{Q}_3 =& P_1\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_2 + P_1\mathbf{G}_2^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_1 \\ &+ P_1\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_1 + P_1\mathbf{G}_2^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_2; \end{aligned} \tag{5.68}$$

$$\mathbf{q}_3^\dagger = P_1(\bar{\mathbf{h}}\odot\check{\mathbf{g}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_2 + P_1(\bar{\mathbf{h}}\odot\check{\mathbf{g}})^T\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_1; \tag{5.69}$$

$$k_3 = P_1(\bar{\mathbf{h}}\odot\check{\mathbf{g}})^T\bar{\mathbf{W}}(\bar{\mathbf{h}}\odot\check{\mathbf{g}})^*. \tag{5.70}$$

The numerator of SNR_2, denoted as a_2, can also be written as

$$\begin{aligned}
a_2 =&P_1|\hat{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}}+\tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}}+\hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}}+\tilde{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}}|^2\\
=&P_1|(\bar{\mathbf{h}}\odot\check{\mathbf{g}}+\Delta\bar{\mathbf{h}}\odot\check{\mathbf{g}}+\bar{\mathbf{h}}\odot\Delta\check{\mathbf{g}})^T\mathbf{w}|^2\\
=&P_1|(\bar{\mathbf{h}}\odot\check{\mathbf{g}}+\mathsf{diag}(\check{\mathbf{g}})\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}+\mathsf{diag}(\bar{\mathbf{h}})\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})^T\mathbf{w}|^2\\
=&P1\Bigg((\bar{\mathbf{h}}\odot\check{\mathbf{g}})^T\bar{\mathbf{W}}(\bar{\mathbf{h}}\odot\check{\mathbf{g}})^*\\
&+\Delta\mathbf{c}^T\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}+\Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c}\\
&+\Delta\mathbf{c}^T\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c}\\
&+\Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}+2(\bar{\mathbf{h}}\odot\check{\mathbf{g}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}^*\\
&+2(\bar{\mathbf{h}}\odot\check{\mathbf{g}})^T\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c}^*\Bigg)\\
=&\Delta\mathbf{c}^T\mathbf{Q}_3\Delta\mathbf{c}^*+2\mathtt{Re}(\mathbf{q}_3^\dagger\Delta\mathbf{c}^*)+k_3.
\end{aligned}\tag{5.71}$$

The denominator part of SNR_2, b_2, is then expressed as

$$\begin{aligned}
b_2 =& P_2|\tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{g}}+\hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{g}}|^2+\sigma^2(\|\hat{\mathbf{g}}^T\mathbf{W}+\tilde{\mathbf{g}}^T\mathbf{W}\|^2+1)\\
=& P_2|(\bar{\mathbf{g}}\odot\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})^T\mathbf{w}+(\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}\odot\check{\mathbf{g}})^T\mathbf{w}|^2\\
&+\sigma^2\left(\|\hat{\mathbf{g}}^T\mathbf{W}\|^2+\|\tilde{\mathbf{g}}^T\mathbf{W}\|^2+2\mathtt{Re}(\hat{\mathbf{g}}^T\mathbf{W}\mathbf{W}^{dag}\tilde{\mathbf{g}}^*)+1\right)\\
=&\Delta\mathbf{c}^T\mathbf{Q}_4\Delta\mathbf{c}^*+2\mathtt{Re}(\mathbf{q}_4^\dagger\Delta\mathbf{c}^*)+k_4,
\end{aligned}\tag{5.72}$$

where

$$\begin{aligned}
\mathbf{Q}_4 =&P_2\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_1+P_2\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1\\
&+P_2\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\check{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{g}}^*)\mathbf{D}_R\mathbf{G}_1+P_2\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_1\\
&+\sigma^2\mathbf{G}_1^T\mathbf{D}_L^T\mathbf{E}\odot\bar{\mathbf{W}}\mathbf{D}_L\mathbf{G}_1;
\end{aligned}\tag{5.73}$$

$$\mathbf{q}_4^\dagger =\sigma^2\check{\mathbf{g}}^T\mathbf{E}\odot\bar{\mathbf{W}}\mathbf{D}_L\mathbf{G}_1;\tag{5.74}$$

$$k_4 =\sigma^2(\check{\mathbf{g}}^T\mathbf{E}\odot\bar{\mathbf{W}}\check{\mathbf{g}}^*+1).\tag{5.75}$$

Then, the numerator of SNR_e, denoted as a_e, can be found as

$$\begin{aligned}
a_e =&P_1|(\hat{\mathbf{e}}+\tilde{\mathbf{e}})^T\mathbf{W}(\hat{\mathbf{h}}+\tilde{\mathbf{h}})|^2+P_2|(\hat{\mathbf{e}}+\tilde{\mathbf{e}})^T\mathbf{W}(\hat{\mathbf{g}}+\tilde{\mathbf{g}})|^2+P_1|h_e|^2+P_2|g_e|^2\\
=&P_1|\hat{\mathbf{e}}^T\mathbf{W}\hat{\mathbf{h}}+\tilde{\mathbf{e}}^T\mathbf{W}\hat{\mathbf{h}}+\hat{\mathbf{e}}^T\mathbf{W}\tilde{\mathbf{h}}+\tilde{\mathbf{e}}^T\mathbf{W}\tilde{\mathbf{h}}|^2\\
&+P_2|\hat{\mathbf{e}}^T\mathbf{W}\hat{\mathbf{g}}+\tilde{\mathbf{e}}^T\mathbf{W}\hat{\mathbf{g}}+\hat{\mathbf{e}}^T\mathbf{W}\tilde{\mathbf{g}}+\tilde{\mathbf{e}}^T\mathbf{W}\tilde{\mathbf{g}}|^2+P_1|h_e|^2+P_2|g_e|^2\\
=&P_2|(\bar{\mathbf{g}}\odot\check{\mathbf{e}}+\Delta\bar{\mathbf{g}}\odot\check{\mathbf{e}}+\bar{\mathbf{g}}\odot\Delta\check{\mathbf{e}})^T\mathbf{w}|^2+P_1|(\bar{\mathbf{h}}\odot\check{\mathbf{e}}+\Delta\bar{\mathbf{h}}\odot\check{\mathbf{e}}+\bar{\mathbf{h}}\odot\Delta\check{\mathbf{e}})^T\mathbf{w}|^2+P_1|h_e|^2+P_2|g_e|^2\\
=&P2\Bigg((\bar{\mathbf{g}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}(\bar{\mathbf{g}}\odot\check{\mathbf{e}})^*\\
&+\Delta\mathbf{c}^T\mathbf{G}_3^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_3\Delta\mathbf{c}+\Delta\mathbf{c}^T\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c}\\
&+\Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c}+\Delta\mathbf{c}^T\mathbf{G}_1^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}\\
&+2(\bar{\mathbf{g}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}^*+2(\bar{\mathbf{g}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c}^*\Bigg)\\
&+P1\Bigg((\bar{\mathbf{h}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}(\bar{\mathbf{h}}\odot\check{\mathbf{e}})^*\\
&+\Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\bar{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}+\Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c}\\
&+\Delta\mathbf{c}^T\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c}\\
&+\Delta\mathbf{c}^T\mathbf{G}_2^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{h}}^*)\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}+2(\bar{\mathbf{h}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}^*\Bigg)\\
=&\Delta\mathbf{c}^T\mathbf{Q}_5\Delta\mathbf{c}^*+2\mathtt{Re}(\mathbf{q}_5^\dagger\Delta\mathbf{c}^*)+k_5,
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{Q}_5 =&P_1\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_3+P_1\mathbf{G}_3^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_1\\
&+P_1\mathbf{G}_1^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_1+P_1\mathbf{G}_3^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_3\\
&+P_2\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_3+P_2\mathbf{G}_3^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_2\\
&+P_2\mathbf{G}_2^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_2+P_2\mathbf{G}_3^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{g}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_3; \qquad (5.76)\\
\mathbf{q}_5^\dagger =&P_1(\bar{\mathbf{h}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_2+P_1(\bar{\mathbf{h}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{h}}^*)\mathbf{D}_L\mathbf{G}_1\\
&+P_2(\bar{\mathbf{g}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_2+P_2(\bar{\mathbf{g}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{g}}^*)\mathbf{D}_L\mathbf{G}_1; \qquad (5.77)\\
k_5 =&P_1(\bar{\mathbf{h}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}(\bar{\mathbf{h}}\odot\check{\mathbf{e}})^*+P_2(\bar{\mathbf{g}}\odot\check{\mathbf{e}})^T\bar{\mathbf{W}}(\bar{\mathbf{g}}\odot\check{\mathbf{e}})^*+P_1|h_e|^2+P_2|g_e|^2. \qquad (5.78)
\end{aligned}$$

In addition, the denominator part of SNR_e, denoted as b_e, can be obtained as

$$\begin{aligned}
b_e &= \sigma^2(\|\hat{\mathbf{e}}+\tilde{\mathbf{e}})^T\mathbf{W}\|^2+2)\\
&= \sigma^2\left(\|\hat{\mathbf{e}}^T\mathbf{W}\|^2+|\tilde{\mathbf{e}}^T\mathbf{W}\|^2+2\mathtt{Re}(\hat{\mathbf{e}}^T\mathbf{W}\mathbf{W}^\dagger\tilde{\mathbf{e}}^*)+2\right)\\
&= \Delta\mathbf{c}^T\mathbf{Q}_6\Delta\mathbf{c}^*+2\mathtt{Re}(\mathbf{q}_6^\dagger\Delta\mathbf{c}^*)+k_6, \qquad (5.79)
\end{aligned}$$

where

$$\begin{aligned}\mathbf{Q}_6 =&\mathbf{G}_3^T\mathbf{D}_R^T\mathsf{diag}(\check{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_3+\mathbf{G}_3^T\mathbf{D}_L^T\mathsf{diag}(\bar{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_L\mathbf{G}_3\\&+\mathbf{G}_3^T\mathbf{D}_L^T\mathsf{diag}(\check{\mathbf{h}})\bar{\mathbf{W}}\mathsf{diag}(\check{\mathbf{e}}^*)\mathbf{D}_R\mathbf{G}_3+\mathbf{G}_3^T\mathbf{D}_R^T\mathsf{diag}(\bar{\mathbf{e}})\bar{\mathbf{W}}\mathsf{diag}(\bar{\mathbf{e}}^*)\mathbf{D}_L\mathbf{G}_3\\&+\sigma^2\mathbf{G}_3^T\mathbf{D}_L^T\mathbf{E}\odot\bar{\mathbf{W}}\mathbf{D}_L\mathbf{G}_3;\end{aligned}\tag{5.80}$$

$$\mathbf{q}_6^\dagger =\sigma^2\check{\mathbf{e}}^T\mathbf{E}\odot\bar{\mathbf{W}}\mathbf{D}_L\mathbf{G}_3;\tag{5.81}$$

$$k_6 =\sigma^2(\check{\mathbf{e}}^T\mathbf{E}\odot\bar{\mathbf{W}}\check{\mathbf{e}}^*+1).\tag{5.82}$$

For the relay power's constraint. First generate matrix $\mathbf{K}$ as the commutation matrix such that $\mathsf{vec}(\bar{\mathbf{W}}^T)=\mathbf{K}\mathsf{vec}(\bar{\mathbf{W}})$, we can define

$$\mathbf{Q}_R=P_1\mathbf{G}_1^T\mathbf{D}_L^T\mathbf{E}\odot(\mathbf{K}\bar{\mathbf{W}}\mathbf{K}^T)\mathbf{D}_L\mathbf{G}_1+P_2\mathbf{G}_2^T\mathbf{D}_L^T\mathbf{E}\odot(\mathbf{K}\bar{\mathbf{W}}\mathbf{K}^T)\mathbf{D}_L\mathbf{G}_2;\tag{5.83}$$

$$\mathbf{q}_R^\dagger=P_1\check{\mathbf{h}}\mathbf{E}\odot(\mathbf{K}\bar{\mathbf{W}}\mathbf{K}^T)\mathbf{D}_L\mathbf{G}_1+P_2\check{\mathbf{g}}\mathbf{E}\odot(\mathbf{K}\bar{\mathbf{W}}\mathbf{K}^T)\mathbf{D}_L\mathbf{G}_2;\tag{5.84}$$

$$k_R=P_1\check{\mathbf{h}}\mathbf{E}\odot(\mathbf{K}\bar{\mathbf{W}}\mathbf{K}^T)\bar{\mathbf{h}}^*+P_2\check{\mathbf{g}}\mathbf{E}\odot(\mathbf{K}\bar{\mathbf{W}}\mathbf{K}^T)\check{\mathbf{g}}^*+\mathsf{trace}(\bar{\mathbf{W}}),\tag{5.85}$$

which allows us to write

$$P_R=\Delta\mathbf{c}^T\mathbf{Q}_R\Delta\mathbf{c}^*+2\mathtt{Re}(\mathbf{q}_R^\dagger\Delta\mathbf{c}^*)+k_R.\tag{5.86}$$

Now, we go back to our optimization problem in (5.27), and define $\tilde{\Gamma}_1=2^{2\Gamma_1}-1$, $\tilde{\Gamma}_2=2^{2\Gamma_2}-1$ and $\tilde{\Gamma}_3=2^{2\Gamma_3}-1$. Therefore, the secrecy rate based optimization problem can be rewritten as

$$\begin{aligned}&\max_{\mathbf{W}}\ \Gamma_1+\Gamma_2-\Gamma_3\\&\text{s.t.}\ \begin{cases}\mathsf{SNR}_1\geq\tilde{\Gamma_1},\\ \mathsf{SNR}_2\geq\tilde{\Gamma_2},\\ \mathsf{SNR}_e\leq\tilde{\Gamma_3},\\ \quad P_T\geq P_1\|\mathbf{W}\mathbf{h}\|^2+P_2\|\mathbf{W}\mathbf{g}\|^2+\mathsf{trace}(\mathbf{W}\mathbf{W}^\dagger)\sigma^2.\end{cases}\end{aligned}\tag{5.87}$$

Considering the ICSI, the first constraint in (5.87) for SNR_1 is equivalent to

$$\begin{cases}\Delta\mathbf{c}^T(\mathbf{Q}_1-\tilde{\Gamma}_1\mathbf{Q}_2)\Delta\mathbf{c}+2\mathtt{Re}\left((\mathbf{q}_1-\tilde{\Gamma}_1\mathbf{q}_2)^\dagger\Delta\mathbf{c}^*\right)+k_1-\tilde{\Gamma}_1k_2\geq 0,\\ \qquad\qquad\qquad\qquad\qquad\qquad\Delta\mathbf{c}^T\mathbf{G}_1^\dagger\mathbf{G}_1\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{h}}^2,\\ \qquad\qquad\qquad\qquad\qquad\qquad\Delta\mathbf{c}^T\mathbf{G}_2^\dagger\mathbf{G}_2\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{g}}^2.\end{cases}\tag{5.88}$$

Furthermore, the second constraint in (5.87) can be expressed in a similar way as

$$\begin{cases}\Delta\mathbf{c}^T(\mathbf{Q}_3-\tilde{\Gamma}_1\mathbf{Q}_4)\Delta\mathbf{c}+2\mathtt{Re}\left((\mathbf{q}_3-\tilde{\Gamma}_1\mathbf{q}_4)^\dagger\Delta\mathbf{c}^*\right)+k_3-\tilde{\Gamma}_1k_4\geq 0,\\ \qquad\qquad\qquad\qquad\qquad\qquad\Delta\mathbf{c}^T\mathbf{G}_1^\dagger\mathbf{G}_1\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{h}}^2,\\ \qquad\qquad\qquad\qquad\qquad\qquad\Delta\mathbf{c}^T\mathbf{G}_2^\dagger\mathbf{G}_2\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{g}}^2.\end{cases}\tag{5.89}$$

For the third constraint in (5.87), we also have

$$\begin{cases} \Delta\mathbf{c}^T(\mathbf{Q}_5-\tilde{\Gamma}_1\mathbf{Q}_6)\Delta\mathbf{c}+2\mathtt{Re}\left((\mathbf{q}_5-\tilde{\Gamma}_1\mathbf{q}_6)^\dagger\Delta\mathbf{c}^*\right)+k_5-\tilde{\Gamma}_1k_5\geq 0, \\ \Delta\mathbf{c}^T\mathbf{G}_1^\dagger\mathbf{G}_1\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{h}}^2, \\ \Delta\mathbf{c}^T\mathbf{G}_2^\dagger\mathbf{G}_2\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{g}}^2, \\ \Delta\mathbf{c}^T\mathbf{G}_3^\dagger\mathbf{G}_3\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{e}}^2. \end{cases} \tag{5.90}$$

Finally, we can write the replay power constraint as

$$\begin{cases} -\Delta\mathbf{c}^T\mathbf{Q}_R\Delta\mathbf{c}^*-2\mathtt{Re}(\mathbf{q}_R^\dagger\Delta\mathbf{c}^*)-k_R+P_T\geq 0, \\ \Delta\mathbf{c}^T\mathbf{G}_1^\dagger\mathbf{G}_1\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{h}}^2, \\ \Delta\mathbf{c}^T\mathbf{G}_2^\dagger\mathbf{G}_2\Delta\mathbf{c}^*\leq\varepsilon_{\mathbf{g}}^2. \end{cases} \tag{5.91}$$

Here, we use S-lemma to continue our transformation of the optimization problem. Using Lemma 1 and Lemma 2, as presented in Chapter 3, we can get the following LMIs:

$$\mathbf{E}_1=\begin{pmatrix} \mathbf{Q}_1-\tilde{\Gamma}_1\mathbf{Q}_2+\lambda_1\mathbf{G}_1^\dagger\mathbf{G}_1+\lambda_2\mathbf{G}_2^\dagger\mathbf{G}_2 & \mathbf{q}_1-\tilde{\Gamma}_1\mathbf{q}_2 \\ \mathbf{q}_1^\dagger-\tilde{\Gamma}_1\mathbf{q}_2^\dagger & k_1-\tilde{\Gamma}_1k_2-\lambda_1\varepsilon_{\mathbf{h}}^2-\lambda_2\varepsilon_{\mathbf{g}}^2 \end{pmatrix}\succeq 0, \tag{5.92}$$

$$\mathbf{E}_2=\begin{pmatrix} \mathbf{Q}_3-\tilde{\Gamma}_2\mathbf{Q}_4+\lambda_3\mathbf{G}_1^\dagger\mathbf{G}_1+\lambda_4\mathbf{G}_2^\dagger\mathbf{G}_2 & \mathbf{q}_3-\tilde{\Gamma}_2\mathbf{q}_4 \\ \mathbf{q}_3^\dagger-\tilde{\Gamma}_2\mathbf{q}_4^\dagger & k_3-\tilde{\Gamma}_2k_4-\lambda_3\varepsilon_{\mathbf{h}}^2-\lambda_4\varepsilon_{\mathbf{g}}^2 \end{pmatrix}\succeq 0, \tag{5.93}$$

$$\mathbf{E}_3=\begin{pmatrix} \mathbf{Q}_5-\tilde{\Gamma}_3\mathbf{Q}_6+\lambda_5\mathbf{G}_1^\dagger\mathbf{G}_1+\lambda_6\mathbf{G}_2^\dagger\mathbf{G}_2+\lambda_7\mathbf{G}_3^\dagger\mathbf{G}_3 & \mathbf{q}_5-\tilde{\Gamma}_3\mathbf{q}_6 \\ \mathbf{q}_5^\dagger-\tilde{\Gamma}_3\mathbf{q}_6^\dagger & k_5-\tilde{\Gamma}_3k_4-\lambda_5\varepsilon_{\mathbf{h}}^2-\lambda_6\varepsilon_{\mathbf{g}}^2-\lambda_7\varepsilon_{\mathbf{e}}^2 \end{pmatrix}\succeq 0, \tag{5.94}$$

$$\mathbf{E}_4=\begin{pmatrix} -\mathbf{Q}_R+\lambda_8\mathbf{G}_1^\dagger\mathbf{G}_1+\lambda_9\mathbf{G}_2^\dagger\mathbf{G}_2 & -\mathbf{q}_R \\ -\mathbf{q}_R^\dagger & P_R-k_R-\lambda_8\varepsilon_{\mathbf{h}}^2-\lambda_9\varepsilon_{\mathbf{g}}^2 \end{pmatrix}\succeq 0. \tag{5.95}$$

As a result, we can rewrite our optimization problem as

$$\max_{\bar{\mathbf{W}},\lambda_i,i=1,\ldots,9}\ \Gamma_1+\Gamma_2-\Gamma_3 \quad \text{s.t.} \begin{cases} \mathbf{E}_1\succeq 0, \\ \mathbf{E}_2\succeq 0, \\ \mathbf{E}_3\succeq 0, \\ \mathbf{E}_4\succeq 0, \\ \mathsf{rank}(\bar{\mathbf{W}})=1. \end{cases} \tag{5.96}$$

Due to the constraint $\mathsf{rank}(\bar{\mathbf{W}}) = 1$, the problem is still not convex. However, if we relax the rank-1 constraint, we obtain the following SDP optimization problem:

$$\max_{\bar{\mathbf{W}},\lambda_i,i=1,\ldots,9} \quad \Gamma_1 + \Gamma_2 - \Gamma_3$$
$$\text{s.t.} \begin{cases} \mathbf{E}_1 \succeq 0, \\ \mathbf{E}_2 \succeq 0, \\ \mathbf{E}_3 \succeq 0, \\ \mathbf{E}_4 \succeq 0. \end{cases} \tag{5.97}$$

The objective is to find the maximum value of a linear function. As SEDUMI [76] only takes the minimisation of an objective function, we rewrite the problem as

$$\min_{\bar{\mathbf{W}},\lambda_i,i=1\ldots9} \quad \Gamma_3 - \Gamma_1 - \Gamma_2$$
$$\text{s.t.} \begin{cases} \mathbf{E}_1 \succeq 0, \\ \mathbf{E}_2 \succeq 0, \\ \mathbf{E}_3 \succeq 0, \\ \mathbf{E}_4 \succeq 0, \end{cases} \tag{5.98}$$

which can be computed as an SDP problem by standard convex optimization tools.

5.3.1 Rank-One Approximation

As (5.98) is an SDP problem, we can use SEDUMI to obtain the optimal solution. However, the optimal solution $\bar{\mathbf{W}}^*$ may not be a rank-one matrix. Recalling from (5.58)–(5.63), we can define

$$\begin{aligned}\mathbf{F}_1 =& P_2(\bar{\mathbf{g}} \odot \check{\mathbf{h}} + \mathsf{diag}(\check{\mathbf{h}})\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c} + \mathsf{diag}(\bar{\mathbf{g}})\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c})(\bar{\mathbf{g}} \odot \check{\mathbf{h}} + \mathsf{diag}(\check{\mathbf{h}})\mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c} + \mathsf{diag}(\bar{\mathbf{g}})\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c})^T \\ & -\Gamma_1\Big((\bar{\mathbf{h}} \odot \mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c} + \mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c} \odot \check{\mathbf{h}})(\bar{\mathbf{h}} \odot \mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c} + \mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c} \odot \check{\mathbf{h}})^T \\ & +\sigma^2(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{h}} + \mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c})(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{h}} + \mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c})^T\Big).\end{aligned}$$

Similarly, from (5.71)–(5.72), we can $\mathbf{F}_2$ as

$$\begin{aligned}\mathbf{F}_2 =& P_1(\bar{\mathbf{h}} \odot \check{\mathbf{g}} + \mathsf{diag}(\check{\mathbf{g}})\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c} + \mathsf{diag}(\bar{\mathbf{h}})\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})(\bar{\mathbf{h}} \odot \check{\mathbf{g}} + \mathsf{diag}(\check{\mathbf{g}})\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c} + \mathsf{diag}(\bar{\mathbf{h}})\mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})^T \\ & -\Gamma_2\Big((\bar{\mathbf{g}} \odot \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c} + \mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c} \odot \check{\mathbf{g}})(\bar{\mathbf{g}} \odot \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c} + \mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c} \odot \check{\mathbf{g}})^T \\ & +\sigma^2(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{g}} + \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{g}} + \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})^T\Big).\end{aligned}$$

Also, from (5.76)–(5.79), we define

$$\begin{aligned}\mathbf{F}_e &= P_2(\bar{\mathbf{g}} \odot \check{\mathbf{e}} + \Delta\bar{\mathbf{g}} \odot \check{\mathbf{e}} + \bar{\mathbf{g}} \odot \Delta\check{\mathbf{e}})(\bar{\mathbf{g}} \odot \check{\mathbf{e}} + \Delta\bar{\mathbf{g}} \odot \check{\mathbf{e}} + \bar{\mathbf{g}} \odot \Delta\check{\mathbf{e}})^T \\ &+ P_1(\bar{\mathbf{h}} \odot \check{\mathbf{e}} + \Delta\bar{\mathbf{h}} \odot \check{\mathbf{e}} + \bar{\mathbf{h}} \odot \Delta\check{\mathbf{e}})(\bar{\mathbf{h}} \odot \check{\mathbf{e}} + \Delta\bar{\mathbf{h}} \odot \check{\mathbf{e}} + \bar{\mathbf{h}} \odot \Delta\check{\mathbf{e}})^T \\ &- \Gamma_3\left(\sigma^2(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{g}} + \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{g}} + \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c})^T\right).\end{aligned}$$

For the relay power, we can define the matrix

$$\mathbf{F}_R = P_1(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{h}})(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{h}})^T + P_2(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{g}})(\mathbf{1}_{M\times 1} \otimes \hat{\mathbf{g}})^T + \sigma^2\mathbf{1}. \tag{5.99}$$

With the above matrix definitions, we can rewrite the constraints of (5.97) into trace functions as

$$\begin{cases} \mathsf{trace}(\mathbf{F}_1\mathbf{W}) \geq 0, \\ \mathsf{trace}(\mathbf{F}_2\mathbf{W}) \geq 0, \\ \mathsf{trace}(\mathbf{F}_e\mathbf{W}) = 0, \\ \mathsf{trace}(\mathbf{F}_R\mathbf{W}) \leq P_R. \end{cases} \tag{5.100}$$

For the case that the optimal solution of (5.98) is not a rank-1 matrix, we provide algorithms to generate a rank-one matrix that satisfy (5.100). The rank-one decomposition method of optimal $\mathbf{W}$ to get a rank-one solution is presented in Appendix II.

5.4 Simulation Results and Analysis

We assumed that the same power has been allocated to nodes S_1 and S_2. For $P_1 = P_2 = [5, 20]$ dB, we simulate 1000 independent channel realisations for each (m, P_R) pair for the SDP solution in (5.98) with the number of antennas being $M = \{2, 4, 6\}$. With the CSI errors $\varepsilon^2_{\mathbf{h}} = \varepsilon^2_{\mathbf{g}} = \varepsilon^2_{\mathbf{e}} \in \{0.1, 0.2, 0.25\}$, we set the relay's power constraint from 0 dB to 35 dB. For the case when $M = 2$, $P_1 = P_2 = 10$ dB, the sum-secrecy rate results versus the relay power constraint are provided in Fig. 5.1. Similar results are provided in Figs. 5.2–5.5, for the settings, $(M = 4, P_1 = P_2 = 5$ dB), $(M = 4, P_1 = P_2 = 5$ dB), $(M = 4, P_1 = P_2 = 10$ dB), $(M = 4, P_1 = P_2 = 20$ dB), $(M = 6, P_1 = P_2 = 10$ dB). Results illustrate that generally if the relay power increases, the secrecy rate will increase but at some point it will go flat because it is limited by the transmit power of the source nodes. Also, we can see a clear performance separation between robust and non-robust approaches, which shows that the proposed robust optimization is important.

As we see from Fig.5.1, for the scenario that the relay has 2 MIMO antennas and the transmission power $P_1 = P_2 = 10$ dB, the optimal secrecy rate has a similar trend, with the error bound of CSI increasing from 0 to 0.25. The simulation results show the robust beamforming scheme of the relay always out-performs the non robust scheme. With the increase of the relay power from 5 dB to 15 dB, the maximum secrecy rate increases almost linearly with the relay power. For the case that relay power is above 15 dB, the maximum secrecy rate almost stays the same. The maximum power relay needed to

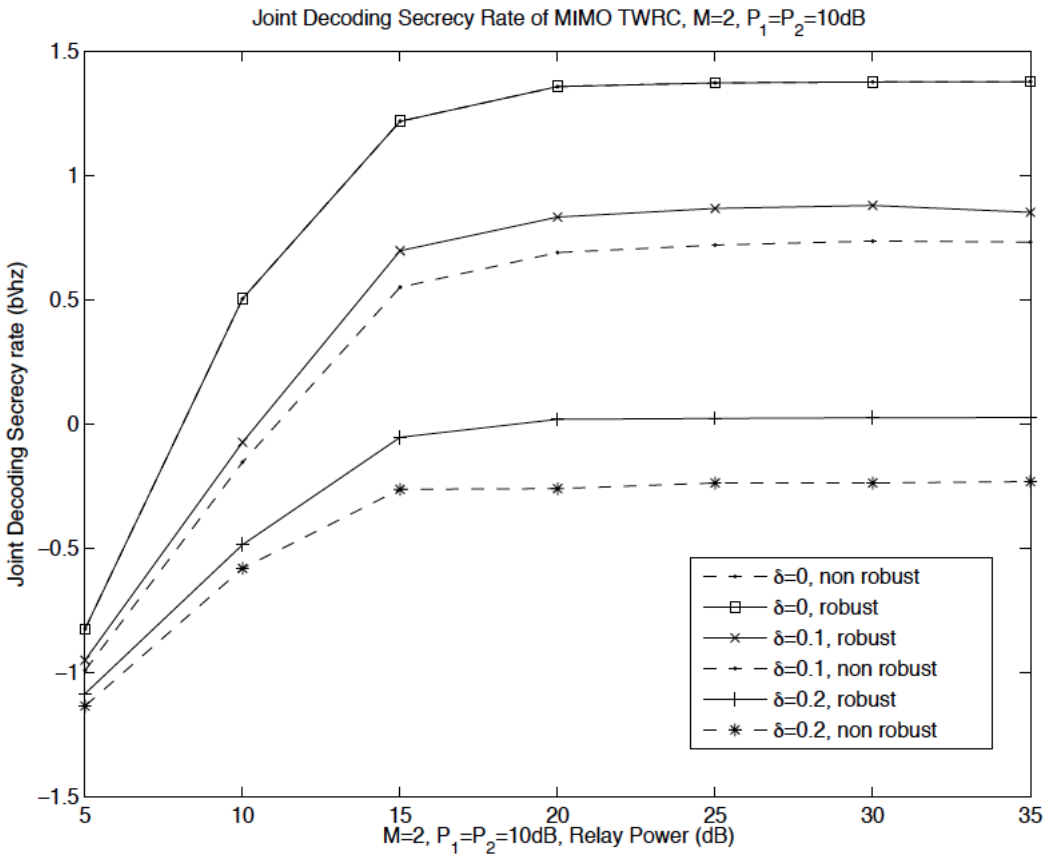

Figure 5.1: Joint decoding secrecy rate of MIMO TWRC, $M = 2$, $P_1 = P_2 = 10$ dB.

maximise the secrecy rate is proportional to the power of P_1 and P_2. The maximum secrecy rate scales with the P_1 and P_2 as well.

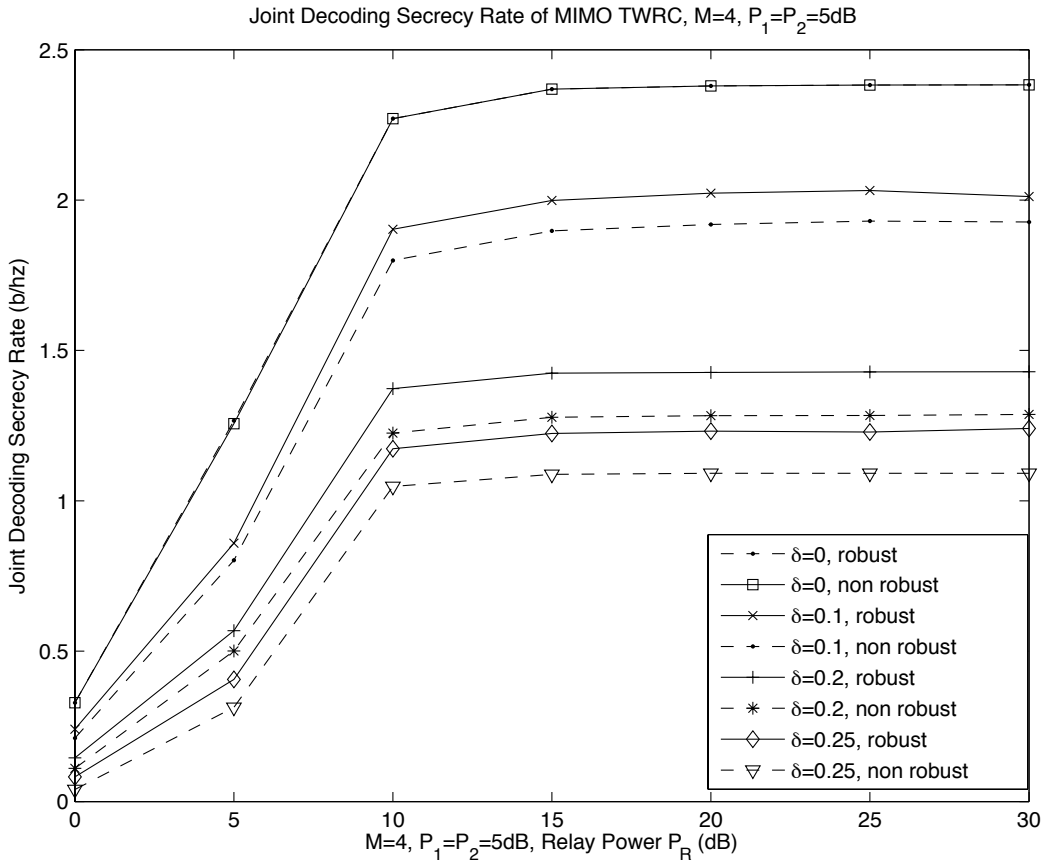

Figure 5.2: Joint decoding secrecy rate of MIMO TWRC, $M = 4$, $P_1 = P_2 = 5$ dB.

Fig.5.2 to Fig.5.4 show the maximal achievable secrecy rates of the scenarios that the relay has 4 MIMO antennas and the transmission power increases from $P_1 = P_2 = 5$ to $P_1 = P_2 = 20$ dB, and with the error bound of CSI increasing from 0 to 0.25. The larger the gap we have of the CSI error, the more reliable the robust beamforming optimization scheme is, compared to the non robust scheme. We also can see that with the same relay power constraint, the maximal achievable secrecy rate increases with the increase of the power of P_1 and P_2, which is because the SNRs of S_1 and S_2 increase lineally with the increase of P_1 and P_2. Also, the minimal power of the relay for the maximum secrecy rate is proportional to the power of P_1 and P_2.

Fig.5.5 shows the maximal achievable secrecy rate of the scenario that the relay has 6 MIMO antennas and the transmission power $P_1 = P_2 = 10$ dB, the optimal secrecy rate has the similar trend, secrecy decreases with the with the error bound of CSI increasing from 0 to 0.25. We can see that the robust beamforming scheme of the relay always out-performs the non robust scheme. With the increase of the relay power from 5 dB to 20 dB, the maximum secrecy rate increases almost linearly with the relay power and becomes stable above that value. As we observe from $M = 4$ cases that there's a minimal usage of relay power to maximise secrecy rate and the minimal relay power is proportional to

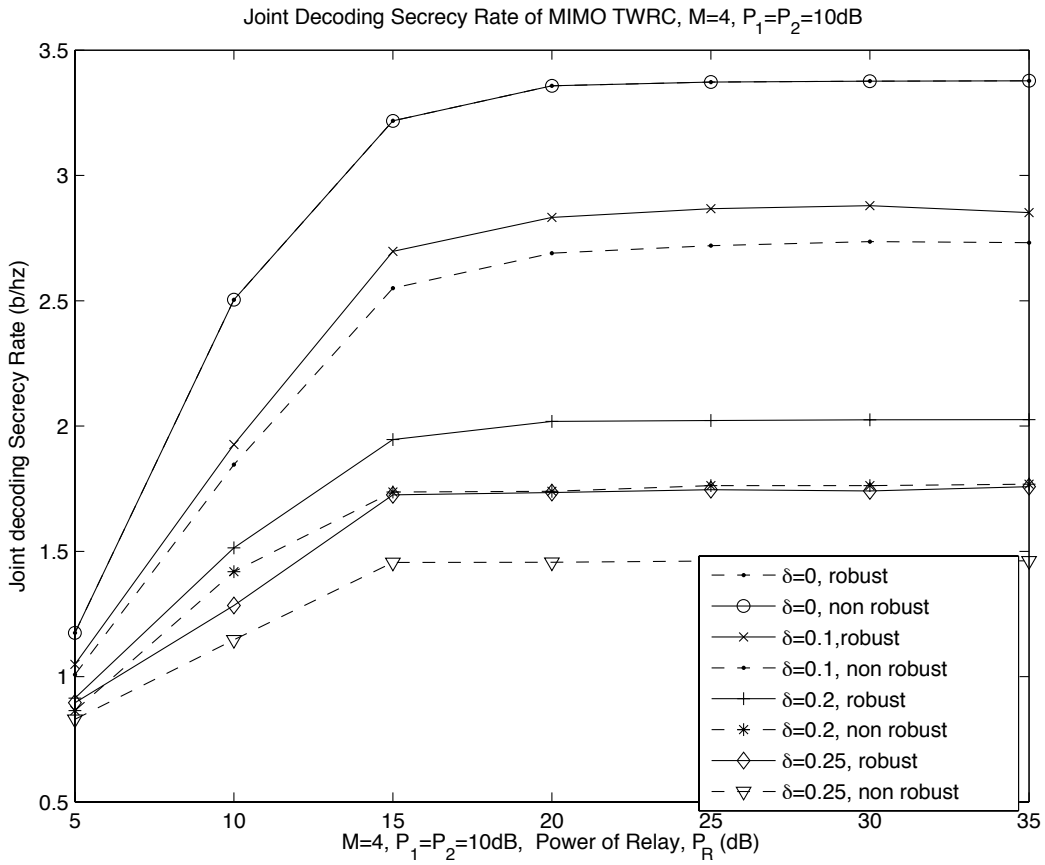

Figure 5.3: Joint decoding secrecy rate of MIMO TWRC, $M = 4$, $P_1 = P_2 = 10$ dB.

the power of P_1 and P_2.

For $P_1 = P_2 = 10$ dB cases with different antenna numbers of the relay ($M = 2, 4, 6$), the simulation shows the maximum sum-secrecy rate increases with the increase of the number of the relay's antenna. Because the more antennas the relay has, the more elements it can adjust from the relay's beamforming matrix to maximize the sum-secrecy rate.

Finally, we investigate the probability that the optimization problem in (5.98) is infeasible or the optimal sum rate is non-positive. For $M = 4$, $P_1 = P_2 = 10$ dB, with $\varepsilon = \{0.1, 0.2\}$, we simulate the outage probability of relay power constraint. The results are shown in Fig. 5.6. It is found that the total relay transmit power required to achieve the positive sum-secrecy rate increases with increase in the CSI error norm bound. The robust beamformer design problem becomes infeasible for the relay power beyond a threshold. From the results, we can observe that this threshold increases with increase in the error norm bound.

5.5 Conclusion

In this chapter, robust beamforming for two-way AF relaying was studied under secrecy constraints. The relay of this TWRC is a MIMO relay which performs beamforming function through its beamforming matrix $\mathbf{W}$. We considered the strict secrecy rate definition of TWRC, which assumes that the eavesdrop-

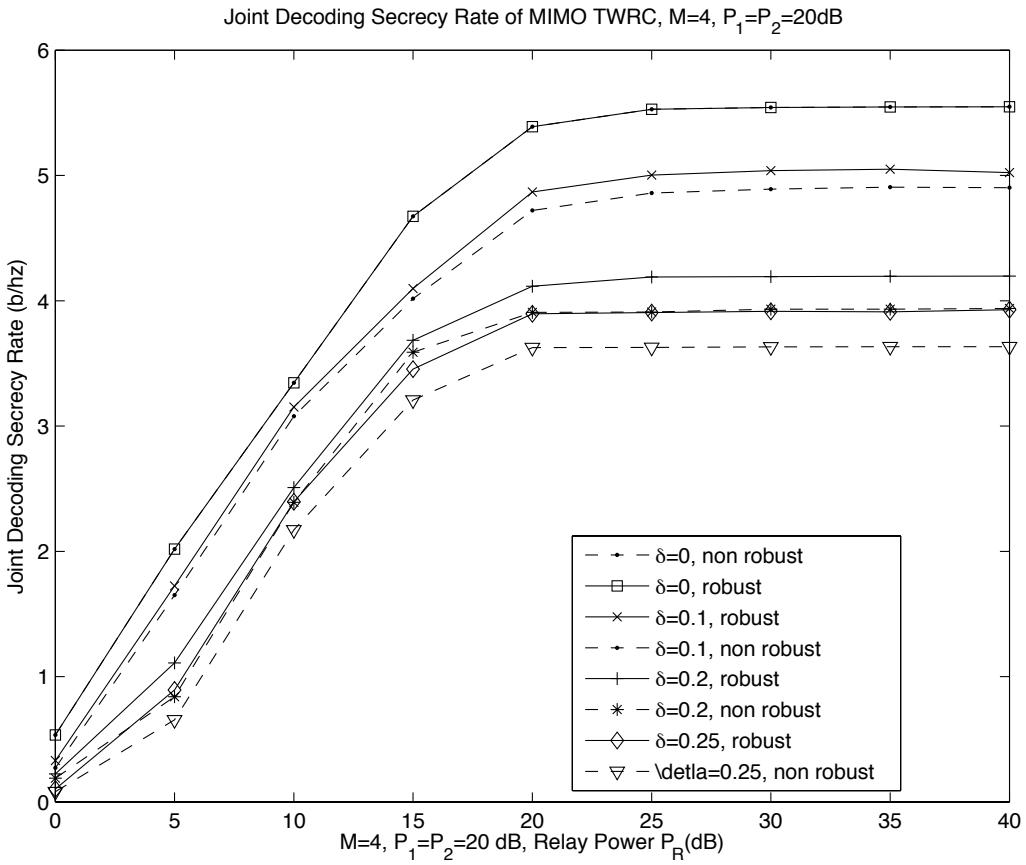

Figure 5.4: Joint decoding secrecy rate of MIMO TWRC, $M = 4$, $P_1 = P_2 = 20$ dB.

per is able to jointly decode the message received. Under these assumptions, we considered to optimize the network to achieve the maximal joint-decoding sum-secrecy rate of S_1 and S_2.

The optimization problem we solved is with an objective function to maximize the joint-decoding secrecy rate of the network and the relay power constraint. ICSI is considered in this chapter, we use an ellipsoid norm bound to describe the ICSI model. And we assumed that the allocation of the power to S_1 and S_2 are the same. For the original non-convex optimization problem, we have presented an SDP formulation, under which optimal beamforming designs that maximize secrecy rates are provided under the relay power constraints, if imperfect CSI is exploited.

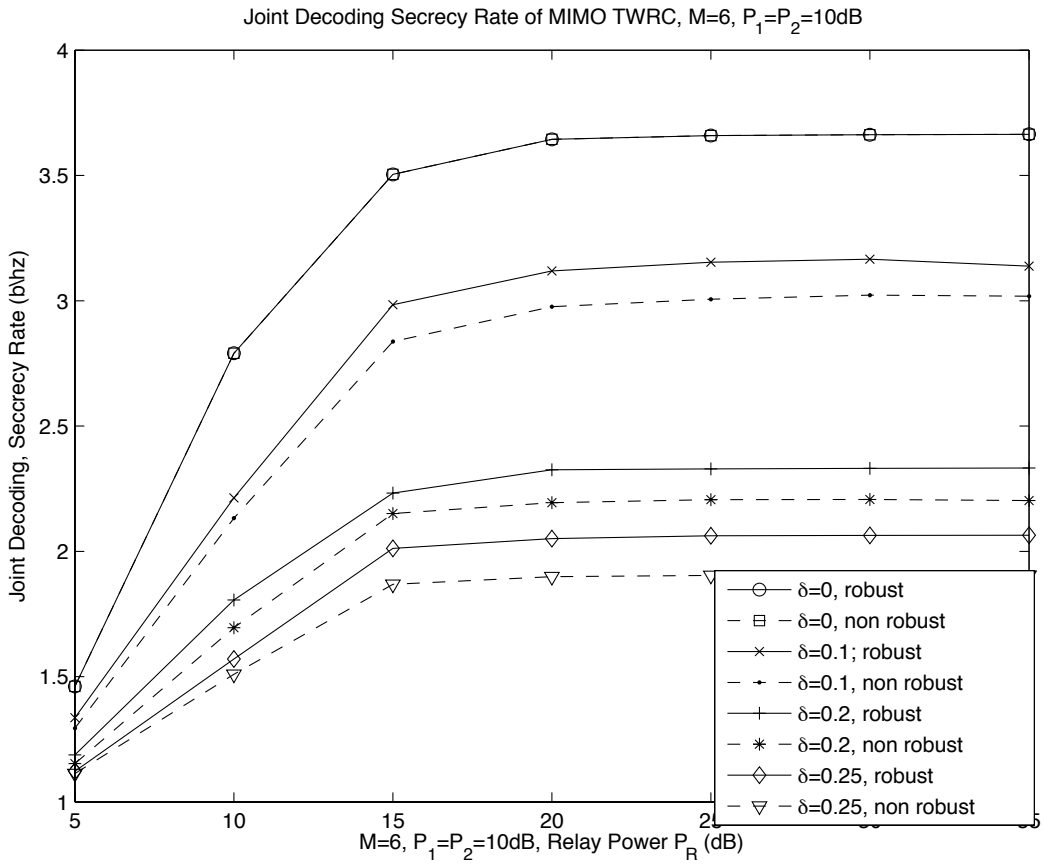

Figure 5.5: Joint decoding secrecy rate of MIMO TWRC, $M = 6$, $P_1 = P_2 = 10$ dB.

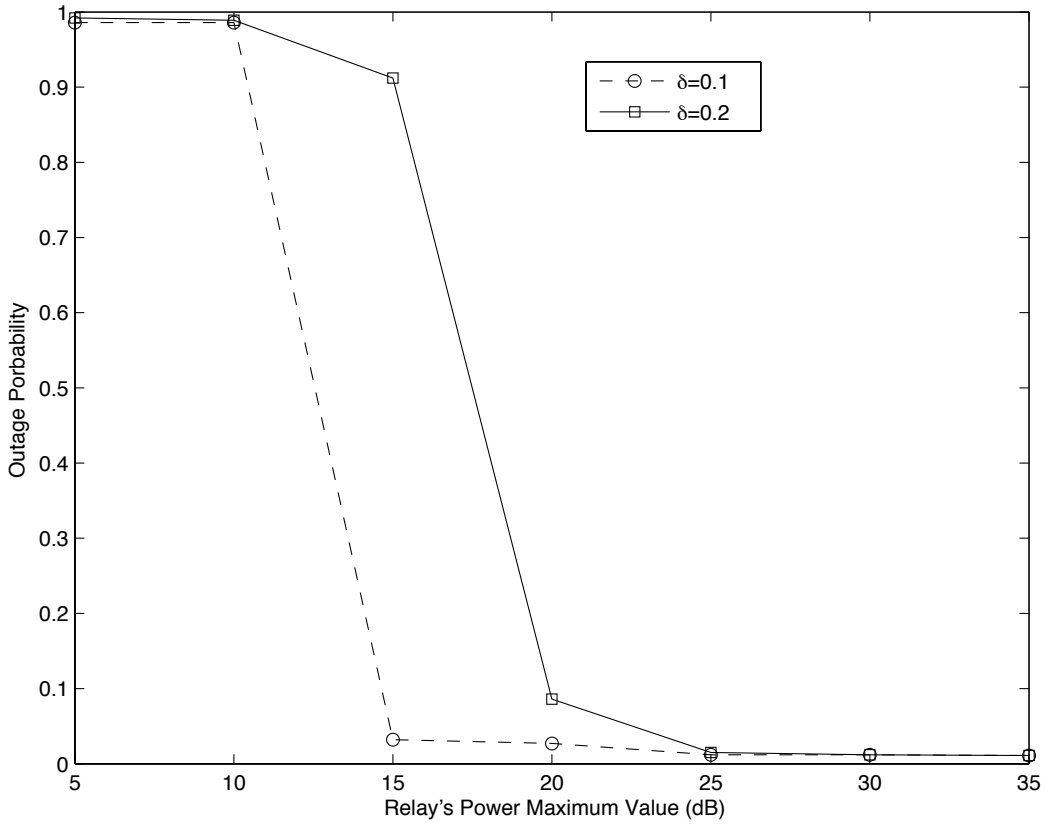

Figure 5.6: Transmit relay power outage probability versus the sum secrecy rate.

Chapter 6

Robust TWRC Beamforming with Untrusted Relay

In this chapter, we revisit the optimization problem of TWBF as in Chapter 6 but instead of having an external eavesdropper, the MIMO relay is considered to be an eavesdropper at the same time, which the message is to be kept confidential from. The formulation is quite similar and our contribution is again that we are able to formulate the optimization problem into an SDP with LMIs and using a rank relaxation method can be solved optimally.

6.1 Network Model

Consider the same TWRC network as before except in the absence of an external eavesdropper. Communications between S_1 and S_2 still takes place in two consecutive time slots but this time the messages are supposed to be kept confidential to the relay. The network model is shown as Fig. 6.1.

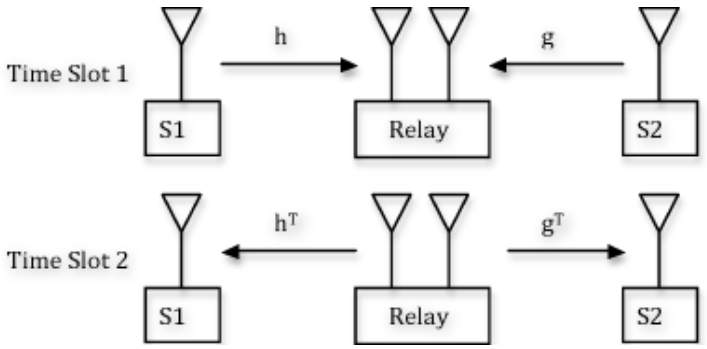

Figure 6.1: The TWRC transmission with an untrusted relay.

During the first time slot, both S_1 and S_2 simultaneously transmit their messages to R. The signals received at R can be represented in vector form as

$$\mathbf{x} = \sqrt{P_1}\mathbf{h}s_1 + \sqrt{P_2}\mathbf{g}s_2 + \mathbf{v}, \tag{6.1}$$

in which P_1 and P_2 are the respective transmit power of S_1 and S_2, s_1 and s_2 denote the symbols transmitted by S_1 and S_2, respectively, and $\mathbf{v} \in \mathcal{C}^M$ is the Gaussian noise vector at R and $\mathbb{E}[\mathbf{v}\mathbf{v}^\dagger] = \sigma^2\mathbf{I}$.

The vectors $\mathbf{h} \in \mathcal{C}^M$, $\mathbf{g} \in \mathcal{C}^M$, are the forward channels from the source nodes to the untrusted relay.

At the second time slot, R transmits a beamformed version of $\mathbf{x}$ by a complex weight matrix $\mathbf{W} \in \mathbb{C}^{M\times M}$, $\mathbf{Wx}$, back to S_1 and S_2. As such, we have the received signals

$$y_1 = \mathbf{h}^T\mathbf{W}\mathbf{x} + \eta_1 \text{ (at } \mathsf{S}_1), \tag{6.2}$$

$$y_2 = \mathbf{g}^T\mathbf{W}\mathbf{x} + \eta_2 \text{ (at } \mathsf{S}_2), \tag{6.3}$$

where η_1, and η_2 denote the respective noise at S_1, and S_2 and they are assumed to be i.i.d. with zero mean and variance of σ^2. In our model, we have assumed that the backward channels from R to S_1 and S_2 are the same as the respective forward channels and they remain static over the period of optimization of interest. Our CSI error model is the same as before, i.e.,

$$\mathbf{h} = \hat{\mathbf{h}} + \tilde{\mathbf{h}}, \tag{6.4}$$

$$\mathbf{g} = \hat{\mathbf{g}} + \tilde{\mathbf{g}}, \tag{6.5}$$

where $\mathbf{h}$ and $\mathbf{g}$ are the true CSI, $\hat{\mathbf{h}}$, $\hat{\mathbf{g}}$ are the imperfect CSI available at the relay nodes, and $\tilde{\mathbf{h}}$, $\tilde{\mathbf{g}}$ represent the additive errors in the CSI. Further, we assume that $\|\tilde{\mathbf{h}}\| \leq \varepsilon_{\mathbf{h}}$, $\|\tilde{\mathbf{g}}\| \leq \varepsilon_{\mathbf{g}}$. Equivalently, $\mathbf{h}$ belongs to the uncertainty set $\mathscr{R}_h$, and $\mathbf{g}$ belongs to the uncertainty set $\mathscr{R}_g$, where

$$\mathscr{R}_h = \left\{\zeta | \zeta = \hat{\mathbf{h}} + \tilde{\mathbf{h}}, \|\tilde{\mathbf{h}}\| \leq \varepsilon_{\mathbf{h}}\right\}, \tag{6.6}$$

$$\mathscr{R}_g = \left\{\zeta | \zeta = \hat{\mathbf{g}} + \tilde{\mathbf{g}}, \|\tilde{\mathbf{g}}\| \leq \varepsilon_{\mathbf{g}}\right\}. \tag{6.7}$$

Take the imperfect CSI model into (6.2)-(6.3), we have

$$\begin{aligned} y_1 =& \sqrt{P_1}(\hat{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{h}} + \tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{h}} + \tilde{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{h}})s_1 \\ &+ \sqrt{P_2}(\hat{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}} + \tilde{\mathbf{h}}^T\mathbf{W}\tilde{\mathbf{g}})s_2 + (\hat{\mathbf{h}}^T\mathbf{W} + \tilde{\mathbf{h}}^T\mathbf{W})\mathbf{v} + \eta_1, \end{aligned} \tag{6.8}$$

$$\begin{aligned} y_2 =& \sqrt{P_1}(\hat{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{h}} + \hat{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}} + \tilde{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{h}})s_1 \\ &+ \sqrt{P_2}(\hat{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{g}} + \tilde{\mathbf{g}}^T\mathbf{W}\hat{\mathbf{g}} + \hat{\mathbf{g}}^T = \mathbf{W}\tilde{\mathbf{g}} + \tilde{\mathbf{g}}^T\mathbf{W}\tilde{\mathbf{g}})s_2 + (\hat{\mathbf{g}}^T\mathbf{W} + \tilde{\mathbf{g}}^T\mathbf{W})\mathbf{v} + \eta_2. \end{aligned} \tag{6.9}$$

6.2 Optimal Beamforming

As before, we assume that the second order CSI error can be neglected. Therefore, after self interference cancellation, the perfect secrecy rates at S_1 and S_2 are given, respectively, as

$$R_1 = \frac{1}{2}\log\left(1 + \mathsf{SNR}_1\right) - \frac{1}{2}\log\left(1 + \mathsf{SNR}_{e1}\right), \tag{6.10}$$

$$R_2 = \frac{1}{2}\log\left(1 + \mathsf{SNR}_2\right) - \frac{1}{2}\log\left(1 + \mathsf{SNR}_{e2}\right). \tag{6.11}$$

We assume that the untrusted relay does the joint decoding for s_1 and s_2. Therefore, the rate based on joint decoding is given by

$$R_r = \frac{1}{2}\log\left(1 + \frac{P_1\|\mathbf{h}\|^2 + P_1\|\mathbf{g}\|^2}{\sigma^2}\right) \tag{6.12}$$

and therefore, we get

$$\begin{aligned} R_s &= R_1 + R_2 - R_r \\ &= \frac{1}{2}\log\left(1+\mathsf{SNR}_1\right) + \frac{1}{2}\log\left(1+\mathsf{SNR}_2\right) - \frac{1}{2}\log\left(1+\mathsf{SNR}_r\right), \end{aligned} \tag{6.13}$$

where

$$\mathsf{SNR}_1 = \frac{P_2|\hat{\mathbf{h}}\mathbf{W}\hat{\mathbf{g}}^T + \tilde{\mathbf{h}}\mathbf{W}\hat{\mathbf{g}}^T + \hat{\mathbf{h}}\mathbf{W}\tilde{\mathbf{g}}^T + \tilde{\mathbf{h}}\mathbf{W}\tilde{\mathbf{g}}^T|^2}{P_1|\tilde{\mathbf{h}}\mathbf{W}\hat{\mathbf{h}}^T + \hat{\mathbf{h}}\mathbf{W}\tilde{\mathbf{h}}^T|^2 + \sigma^2(\|\hat{\mathbf{h}}\mathbf{W} + \tilde{\mathbf{h}}\mathbf{W}\|^2 + 1)}, \tag{6.14}$$

$$\mathsf{SNR}_2 = \frac{P_1|\hat{\mathbf{g}}\mathbf{W}\hat{\mathbf{h}}^T + \tilde{\mathbf{g}}\mathbf{W}\hat{\mathbf{h}}^T + \hat{\mathbf{g}}\mathbf{W}\tilde{\mathbf{h}}^T + \tilde{\mathbf{g}}\mathbf{W}\tilde{\mathbf{h}}^T|^2}{P_2|\tilde{\mathbf{g}}\mathbf{W}\hat{\mathbf{h}}^T + \hat{\mathbf{g}}\mathbf{W}\tilde{\mathbf{g}}^T|^2 + \sigma^2(\|\hat{\mathbf{g}}\mathbf{W} + \tilde{\mathbf{g}}\mathbf{W}\|^2 + 1)}, \tag{6.15}$$

$$\mathsf{SNR}_r = \frac{P_1\|\mathbf{h}\|^2 + P_2\|\mathbf{g}\|^2}{\sigma^2} \equiv \Gamma. \tag{6.16}$$

As a result, the joint decoding based secrecy rate is given by

$$\begin{aligned} R_s =& \frac{1}{2}\log\left\{1 + \frac{P_2|\hat{\mathbf{h}}\mathbf{W}\hat{\mathbf{g}}^T + \tilde{\mathbf{h}}\mathbf{W}\hat{\mathbf{g}}^T + \hat{\mathbf{h}}\mathbf{W}\tilde{\mathbf{g}}^T + \tilde{\mathbf{h}}\mathbf{W}\tilde{\mathbf{g}}^T|^2}{P_1|\tilde{\mathbf{h}}\mathbf{W}\hat{\mathbf{h}}^T + \hat{\mathbf{h}}\mathbf{W}\tilde{\mathbf{h}}^T|^2 + \sigma^2(\|\hat{\mathbf{h}}\mathbf{W} + \tilde{\mathbf{h}}\mathbf{W}\|^2 + 1)}\right\} \\ &+ \frac{1}{2}\log\left\{1 + \frac{P_1|\hat{\mathbf{g}}\mathbf{W}\hat{\mathbf{h}}^T + \tilde{\mathbf{g}}\mathbf{W}\hat{\mathbf{h}}^T + \hat{\mathbf{g}}\mathbf{W}\tilde{\mathbf{h}}^T + \tilde{\mathbf{g}}\mathbf{W}\tilde{\mathbf{h}}^T|^2}{P_2|\tilde{\mathbf{g}}\mathbf{W}\hat{\mathbf{h}}^T + \hat{\mathbf{g}}\mathbf{W}\tilde{\mathbf{g}}^T|^2 + \sigma^2(\|\hat{\mathbf{g}}\mathbf{W} + \tilde{\mathbf{g}}\mathbf{W}\|^2 + 1)}\right\} - \frac{1}{2}\log(1+\Gamma). \end{aligned} \tag{6.17}$$

Our objective is to maximize the sum secrecy rate R_s subject to the relay power constraint, i.e.,

$$\begin{aligned} &\max_{\mathbf{W}} \; R_s \\ &\text{s.t.} \; \Big\{ P_1\|\mathbf{W}\mathbf{h}\|^2 + P_2\|\mathbf{W}\mathbf{g}\|^2 + \mathsf{trace}(\mathbf{W}\mathbf{W}^\dagger)\sigma^2 \le P_T. \end{aligned} \tag{6.18}$$

The above problem is again a non-convex problem. To solve the problem, it is necessary to rewrite the constraints into LMIs. Doing so will require the same definitions as in Section 6.3. In addition, we have slightly different definitions as follows:

$$\mathbf{c} = \mathsf{vec}(\mathbf{h}, \mathbf{g}), \tag{6.19}$$

$$\Delta\mathbf{c} = \mathsf{vec}(\tilde{\mathbf{h}}, \tilde{\mathbf{g}}), \tag{6.20}$$

$$\mathbf{G}_1 = \left[\begin{array}{cc} \mathbf{I}_M & \mathbf{O}_M \end{array}\right] \in \mathcal{R}^{M\times 2M}, \tag{6.21}$$

$$\mathbf{G}_2 = \left[\begin{array}{cc} \mathbf{O}_M & \mathbf{I}_M \end{array}\right] \in \mathcal{R}^{M\times 2M}. \tag{6.22}$$

After these definition, we can easily check that

$$\mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c} = \tilde{\mathbf{h}} \otimes \mathbf{1}_{M\times 1} \tag{6.23}$$

and

$$\mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c} = \mathbf{1}_{M\times 1}\otimes\tilde{\mathbf{h}}. \tag{6.24}$$

Then we define the following vectors:

$$\Delta\bar{\mathbf{h}} = \tilde{\mathbf{h}}\otimes\mathbf{1}_{M\times 1} = \mathbf{D}_R\mathbf{G}_1\Delta\mathbf{c}; \tag{6.25}$$

$$\Delta\check{\mathbf{h}} = \mathbf{1}_{M\times 1}\otimes\tilde{\mathbf{h}} = \mathbf{D}_L\mathbf{G}_1\Delta\mathbf{c}; \tag{6.26}$$

$$\Delta\bar{\mathbf{g}} = \tilde{\mathbf{g}}\otimes\mathbf{1}_{M\times 1} = \mathbf{D}_R\mathbf{G}_2\Delta\mathbf{c}; \tag{6.27}$$

$$\Delta\check{\mathbf{g}} = \mathbf{1}_{M\times 1}\otimes\tilde{\mathbf{g}} = \mathbf{D}_L\mathbf{G}_2\Delta\mathbf{c}. \tag{6.28}$$

For SNR_1, we can, as in Chapter 6, express the signal part as

$$\mathbf{h}^T\mathbf{W}\mathbf{g} = [(\mathbf{g}\otimes\mathbf{1}_{M\times 1})\odot(\mathbf{1}_{M\times 1})\otimes\mathbf{h})]^T\mathsf{vec}(\mathbf{W}). \tag{6.29}$$

Furthermore, we can work out the numerator and the denominator of SNR_1, respectively, in the form:

$$a_1 = \Delta\mathbf{c}^T\mathbf{Q}_1\Delta\mathbf{c}^* + 2\mathtt{Re}(\mathbf{q}_1^\dagger\Delta\mathbf{c}^*) + k_1, \tag{6.30}$$

$$b_1 = \Delta\mathbf{c}^T\mathbf{Q}_2\Delta\mathbf{c}^* + 2\mathtt{Re}(\mathbf{q}_2^\dagger\Delta\mathbf{c}^*) + k_2. \tag{6.31}$$

Also, we can do the same for SNR_2 to have

$$a_2 = \Delta\mathbf{c}^T\mathbf{Q}_3\Delta\mathbf{c}^* + 2\mathtt{Re}(\mathbf{q}_3^\dagger\Delta\mathbf{c}^*) + k_3, \tag{6.32}$$

$$b_2 = \Delta\mathbf{c}^T\mathbf{Q}_4\Delta\mathbf{c}^* + 2\mathtt{Re}(\mathbf{q}_4^\dagger\Delta\mathbf{c}^*) + k_4. \tag{6.33}$$

The relay power constraint can also be rewritten as

$$P_R = \Delta\mathbf{c}^T\mathbf{Q}_R\Delta\mathbf{c}^* + 2\mathtt{Re}(\mathbf{q}_R^\dagger\Delta\mathbf{c}^*) + k_R. \tag{6.34}$$

The secrecy rate based optimization problem can therefore be recast into

$$\begin{aligned}
&\max_{\mathbf{W}}\ \Gamma_1+\Gamma_2-\Gamma_3\\
&\text{s.t.}\ \begin{cases}
\mathsf{SNR}_1 \geq \tilde{\Gamma}_1,\\
\mathsf{SNR}_2 \geq \tilde{\Gamma}_2,\\
\quad P_T \geq P_1\|\mathbf{W}\mathbf{h}\|^2 + P_2\|\mathbf{W}\mathbf{g}\|^2 + \mathsf{trace}(\mathbf{W}\mathbf{W}^\dagger)\sigma^2 P_T.
\end{cases}
\end{aligned} \tag{6.35}$$

Now, managing the ICSI errors in the same way as in Chapter 6, we get the optimization problem

$$\max_{\bar{\mathbf{W}},\lambda_i,i=1,\ldots,6} \quad \Gamma_1+\Gamma_2-\Gamma_3 \qquad \text{s.t.} \begin{cases} \mathbf{E}_1 \succeq 0, \\ \mathbf{E}_2 \succeq 0, \\ \mathbf{E}_3 \succeq 0, \\ \mathsf{rank}(\bar{\mathbf{W}})=1, \end{cases} \tag{6.36}$$

where the LMI constraints are due to the S-Lemmas, or

$$\mathbf{E}_1 = \begin{pmatrix} \mathbf{Q}_1-\Gamma_1\mathbf{Q}_2+\lambda_1\mathbf{G}_1^\dagger\mathbf{G}_1+\lambda_2\mathbf{G}_2^\dagger\mathbf{G}_2 & \mathbf{q}_1-\tilde{\Gamma}_1\mathbf{q}_2 \\ \mathbf{q}_1^\dagger-\tilde{\Gamma}_1\mathbf{q}_2^\dagger & k_1-\tilde{\Gamma}_1k_2-\lambda_1\varepsilon_{\mathbf{h}}^2-\lambda_2\varepsilon_{\mathbf{g}}^2 \end{pmatrix} \succeq 0, \tag{6.37}$$

$$\mathbf{E}_2 = \begin{pmatrix} \mathbf{Q}_3-\Gamma_2\mathbf{Q}_4+\lambda_3\mathbf{G}_1^\dagger\mathbf{G}_1+\lambda_4\mathbf{G}_2^\dagger\mathbf{G}_2 & \mathbf{q}_3-\tilde{\Gamma}_2\mathbf{q}_4 \\ \mathbf{q}_3^\dagger-\tilde{\Gamma}_2\mathbf{q}_4^\dagger & k_3-\tilde{\Gamma}_2k_4-\lambda_3\varepsilon_{\mathbf{h}}^2-\lambda_4\varepsilon_{\mathbf{g}}^2 \end{pmatrix} \succeq 0, \tag{6.38}$$

$$\mathbf{E}_3 = \begin{pmatrix} -\mathbf{Q}_R+\lambda_5\mathbf{G}_1^\dagger\mathbf{G}_1+\lambda_6\mathbf{G}_2^\dagger\mathbf{G}_2 & -\mathbf{q}_R \\ -\mathbf{q}_R^\dagger & P_T-k_R-\lambda_5\varepsilon_{\mathbf{h}}^2-\lambda_6\varepsilon_{\mathbf{g}}^2 \end{pmatrix} \succeq 0. \tag{6.39}$$

Due to the constraint $\mathsf{rank}(\bar{\mathbf{W}})=1$, the problem (6.36) is not convex. However, if we remove the rank-1 constraint, we get the SDP optimization problem:

$$\max_{\bar{\mathbf{W}},\lambda_i,i=1\ldots6} \quad \Gamma_1+\Gamma_2-\Gamma_3 \qquad \text{s.t.} \begin{cases} \mathbf{E}_1 \succeq 0, \\ \mathbf{E}_2 \succeq 0, \\ \mathbf{E}_3 \succeq 0. \end{cases} \tag{6.40}$$

To facilitate optimization using SEDUMI, we rewrite the problem as

$$\min_{\bar{\mathbf{W}},\lambda_i,i=1\ldots9} \quad \Gamma_3-\Gamma_1-\Gamma_2 \qquad \text{s.t.} \begin{cases} \mathbf{E}_1 \succeq 0, \\ \mathbf{E}_2 \succeq 0, \\ \mathbf{E}_3 \succeq 0. \end{cases} \tag{6.41}$$

6.3 Optimal Structure with Perfect CSI

In this section, we present the optimal structure of relay beamforming matrix, assuming that perfect CSI is available for all the nodes within the network. In this case, recall that

$$R_s = \frac{1}{2}\log\left\{1+\frac{P_2|\mathbf{h}^T\mathbf{W}\mathbf{g}|^2}{\sigma^2\left(\|\mathbf{h}^T\mathbf{W}\|^2+1\right)}\right\}+\frac{1}{2}\log\left\{1+\frac{P_1|\mathbf{g}^T\mathbf{W}\mathbf{h}|^2}{\sigma^2\left(\|\mathbf{g}^T\mathbf{W}\|^2+1\right)}\right\}. \tag{6.42}$$

Note that the beamforming matrix only influences the transceivers' rate R_1 and R_2, and has no influence on the joint decoding rate of the untrusted relay node. Therefore, the relay beamforming matrix which maximizes the secrecy rate is actually the same as the one that maximizes the sum-rate of $R_1 + R_2$. Moreover, the optimal beamforming matrix takes the form as (3.19) in Chapter 3: $\mathbf{W} = (\mathbf{U}^{\|})^{*}\mathbf{B}(\mathbf{U}^{\|})^{\dagger}$. As a result, the optimization problem becomes

$$\mathbb{P} \mapsto \begin{cases} \min\limits_{\mathbf{B}} \ \dfrac{1}{2}\log\left\{1 + \dfrac{P_2|\mathbf{h}_1^T\mathbf{B}\mathbf{g}|^2}{\sigma^2\left(\|\mathbf{h}_1^T\mathbf{B}\|^2+1\right)}\right\} + \dfrac{1}{2}\log\left\{1 + \dfrac{P_1|\mathbf{g}_1^T\mathbf{B}\mathbf{h}_1|^2}{\sigma^2\left(\|\mathbf{g}_1^T\mathbf{B}\|^2+1\right)}\right\} \\ \text{s.t.} \ \ P_1\|\mathbf{B}\mathbf{h}_1\|^2 + P_2\|\mathbf{B}\mathbf{g}_1\|^2 + \mathsf{trace}(\mathbf{B}\mathbf{B}^{\dagger})\sigma^2 \le P_T, \end{cases} \tag{6.43}$$

in which $\mathbf{h}_1 \triangleq (\mathbf{U}^{\|})^{\dagger}\mathbf{h} \in \mathcal{C}^{2\times 1}$, $\mathbf{g}_1 \triangleq (\mathbf{U}^{\|})^{\dagger}\mathbf{g} \in \mathcal{C}^{2\times 1}$. The objective function of (6.43) is not convex, although the constraint is convex. Here, we use the traditional weighted sum rate method to look for an upper-bound of the sum-secrecy rate. Let $\beta_1 \ge 0$ and $\beta_2 \ge 0$ be the weighting factors of R_1 and R_2, and express the weighted sum-secrecy rate as

$$R_{\mathsf{WSR}} = \frac{\beta_1}{2}\log\left[1 + \frac{P_2|\mathbf{h}_1^T\mathbf{B}\mathbf{g}_1|^2}{\sigma^2\left(\|\mathbf{h}_1^T\mathbf{B}\|^2+1\right)}\right] + \frac{\beta_2}{2}\log\left[1 + \frac{P_1|\mathbf{g}_1^T\mathbf{B}\mathbf{h}_1|^2}{\sigma^2\left(\|\mathbf{g}_1^T\mathbf{B}\|^2+1\right)}\right]. \tag{6.44}$$

Since (6.44) is still non-convex, we let

$$\alpha_1 = \frac{R_1}{R_s}, \text{ and } \alpha_2 = \frac{R_2}{R_s}. \tag{6.45}$$

For a given $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]^T$, consider the following sum-secrecy-rate optimization problem :

$$\mathbb{P} \mapsto \begin{cases} \min\limits_{\mathbf{B}, R_s} \ R_s \\ \text{s.t.} \ \begin{cases} \dfrac{\beta_1}{2}\log\left\{(1 + \dfrac{P_2|\mathbf{h}^T\mathbf{B}\mathbf{g}|^2}{\sigma^2\left(\|\mathbf{h}^T\mathbf{B}\|^2+1\right)}\right\} \ge \alpha_1 R_s \\ \dfrac{\beta_2}{2}\log\left\{(1 + \dfrac{P_1|\mathbf{g}^T\mathbf{B}\mathbf{h}|^2}{\sigma^2\left(\|\mathbf{g}^T\mathbf{B}\|^2+1\right)}\right\} \ge \alpha_2 R_s \\ P_1\|\mathbf{B}\mathbf{h}\|^2 + P_2\|\mathbf{B}\mathbf{g}\|^2 + \mathsf{trace}(\mathbf{B}\mathbf{B}^{\dagger})\sigma^2 \le P_R. \end{cases} \end{cases} \tag{6.46}$$

To solve (6.46), we can solve the bellowing problem at first:

$$\mathbb{P} \mapsto \begin{cases} \min\limits_{\mathbf{B}} \ P_1\|\mathbf{B}\mathbf{h}_1\|^2 + P_2\|\mathbf{B}\mathbf{g}_1\|^2 + \mathsf{trace}(\mathbf{B}\mathbf{B}^{\dagger})\sigma^2 \\ \text{s.t.} \ \begin{cases} \dfrac{1}{2}\log\left\{(1 + \dfrac{P_2|\mathbf{h}_1^T\mathbf{B}\mathbf{g}_1|^2}{\sigma^2\left(\|\mathbf{h}_1^T\mathbf{B}\|^2+1\right)}\right\} \ge \alpha_1 r, \\ \dfrac{1}{2}\log\left\{(1 + \dfrac{P_1|\mathbf{g}_1^T\mathbf{B}\mathbf{h}_1|^2}{\sigma^2\left(\|\mathbf{g}_1^T\mathbf{B}\|^2+1\right)}\right\} \ge \alpha_2 r. \end{cases} \end{cases} \tag{6.47}$$

For the given $\boldsymbol{\alpha}$ and $r = R_s$, if (6.46) is solvable, then the optimal relay power, denoted as P_R^* , is the minimum relay power that support the given pair $(\boldsymbol{\alpha}, R_s)$, we have the optimal beamforming matrix $\mathbf{W}$ corresponding to the boundary of $\mathbf{R} = (P_1, P_2, P_R)$. Otherwise, there is no finite relay power that

can support $\boldsymbol{\alpha}$ and $r = R_s$ rate pair.

Combining the problems (6.46) and (6.47) together, we have given values of P_R, $\boldsymbol{\alpha}$ and $r = R_s$. If the optimal solution of (6.47) satisfying $P_R^* > P_R$, it means that the R_s obtained in (6.47) is not a feasible solution for (6.46). Otherwise, the value we used for R_s in (6.47) is a feasible point for (6.46), or it is within the sum-secrecy-rate boundary achieved by the optimization problem (6.47). Based on this observation, we use the algorithm below to solve (6.46).

Algorithm 2 Algorithmic to find optimal Sum-secrecy Rate of (6.46)

$R_s \in [0, \check{R}_s]$, α
Initialize $R_{s,min} = 0$,$R_{s,max} = \check{R}_s$;
while $R_{s,max} - R_{s,min} \geq \varepsilon$ **do**
  $R_s = \frac{R_{s,min}+R_{s,max}}{2}$.
  Solve problem (6.47) with the above t_1 and t_2 and get optimal objective value P_R^*
  check the optimal objective solution P_R^* with the P_R in (6.46)
  if $P_R > P_R^*$ **then**
    $R_{s,min} = R_s$
  else
    $R_{s,max} = R_s$
  end if
end while
The $R_{s,min}$ of (6.47) then is the optimal solution of R_s of problem (6.46)

In addition, we also provide a low complexity sub-optimal solution of (6.46), using the dominant eigenvalue method in [19] which is presented in Appendix V.

6.4 Simulation Results

Our simulation results illustrate the optimization results for (6.47). For $P_1 = P_2 = \{5, 10, 20\}$ dB, we simulated 1000 independent channel realisations for each $P_R \in [0, 35]$ dB pair for the SDP solution with antenna number $M = \{2, 4, 6\}$, CSI errors $\varepsilon_{\mathbf{h}}^2 = \varepsilon_{\mathbf{g}}^2 \in \{0, 0.1, 0.2, 0.25\}$. Specifically, Figs. 6.2–6.6 provide similar results but for the settings, respectively, $(M = 2, P_1 = P_2 = 10$ dB), $(M = 4, P_1 = P_2 = 5$ dB), $(M = 4, P_1 = P_2 = 10$ dB), $(M = 2, P_1 = P_2 = 20$ dB), and $(M = 6, P_1 = P_2 = 5$ dB). Results in these figures demonstrate that generally the higher the relay transmit power the better the secrecy rate. However, at some point, it will saturate because of the limit in the source transmit power and the secrecy rate will not get higher. Also, we can see that for larger CSI errors the secrecy rate achievable is lower. Results also clearly show that there is a performance gap between the robust optimal solution and the non-robust one.

As we see from Fig.6.2.For the scenario that the relay has 2 MIMO antennas and the transmission power $P_1 = P_2 = 10$ dB, the optimal secrecy rate decreases, with the error bound of CSI increasing from 0 to 0.25. The simulation result shows the robust beamforming scheme of the relay is always outperform the non robust scheme. With the increase of the relay power from 5 dB to 15 dB, the maximum secrecy rate increases almost linearly with the relay power. But for the increase of relay power above the minimal required relay power, the maximum secrecy rate almost stays the same. The maximum power the relay needs to maximal secrecy rate is proportional to the power of P_1 and P_2.

Fig.6.3 to Fig.6.5 show the maximal achievable secrecy rates of the cases that the relay has 4 MIMO

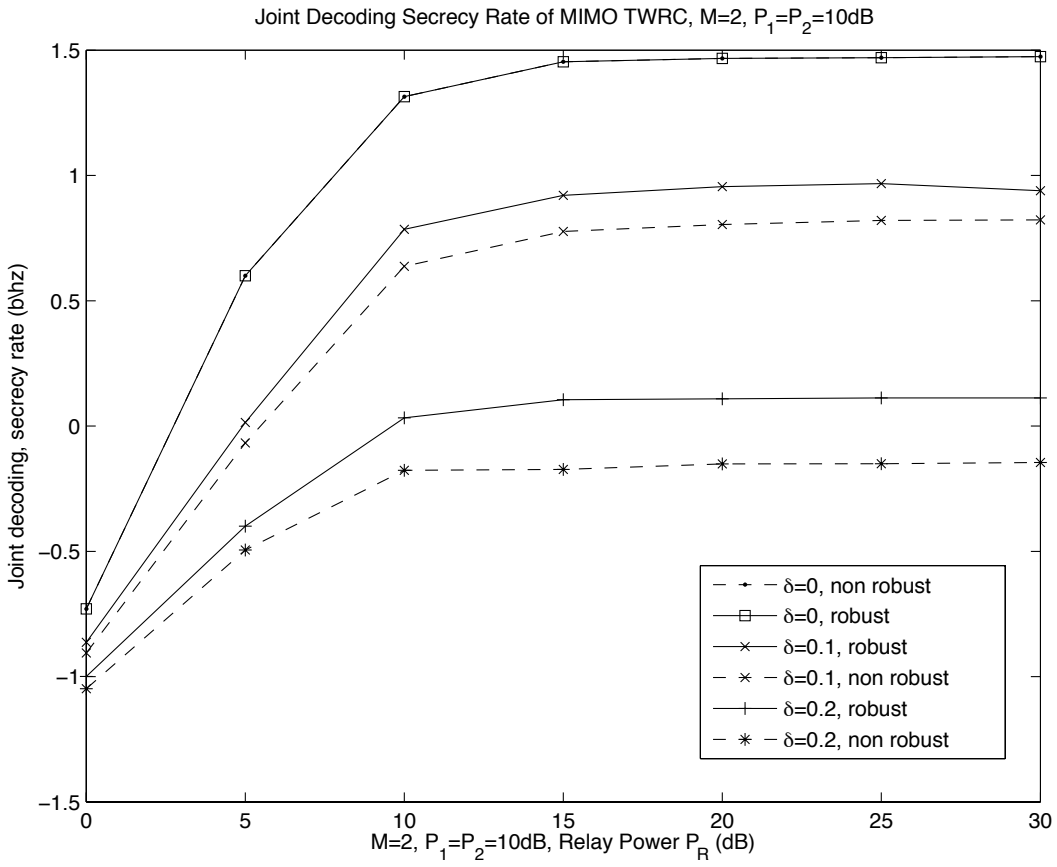

Figure 6.2: Secrecy rate of MIMO TWRC with an untrusted relay, with $M = 2$, $P_1 = P_2 = 10$ dB.

antennas and the transmission power increases from $P_1 = P_2 = 5$ to $P_1 = P_2 = 20$ dB, and with the error bound of CSI increasing from 0 to 0.25. Observing from the simulation results, with the same relay power constraint, the maximal achievable secrecy rate increases with the increase of the power of P_1 and P_2, which is because the SNR of S_1 and S_2 increase proportionally with the increment of P_1 and P_2. Also, the maximum power of the relay needs to achieve maximum secrecy rate scales with the power of P_1 and P_2. From Fig.6.6 we can see that for the scenario that the relay has 6 MIMO antennas and the transmission power $P_1 = P_2 = 10$ dB, the optimal secrecy rate has the similar trend. The maximum secrecy rate decreases with the error bound of CSI increasing from 0 to 0.25. As predicted, robust beamforming scheme of the relay always out-performs the non robust scheme. With the increase of the relay power from 5 dB to 20 dB, the maximum secrecy rate increases almost linearly with the relay power. But when the increase of relay power is above 15 dB, the maximum secrecy rate becomes stable. Because there is a minimal power relay to the maximal secrecy rate and the minimal relay power is proportional to the power of P_1 and P_2. For the case that $P_1 = P_2 = 10db$, simulations results show that the maximum sum-secrecy rate increases with the increase of the number of the relay's antenna. Because the more antennas the relay has, the more elements it can adjust from the its beamforming matrix to maximize the sum-secrecy rate.

Finally, we provide the outage probability results for the optimization problem in Fig. 6.7 and results

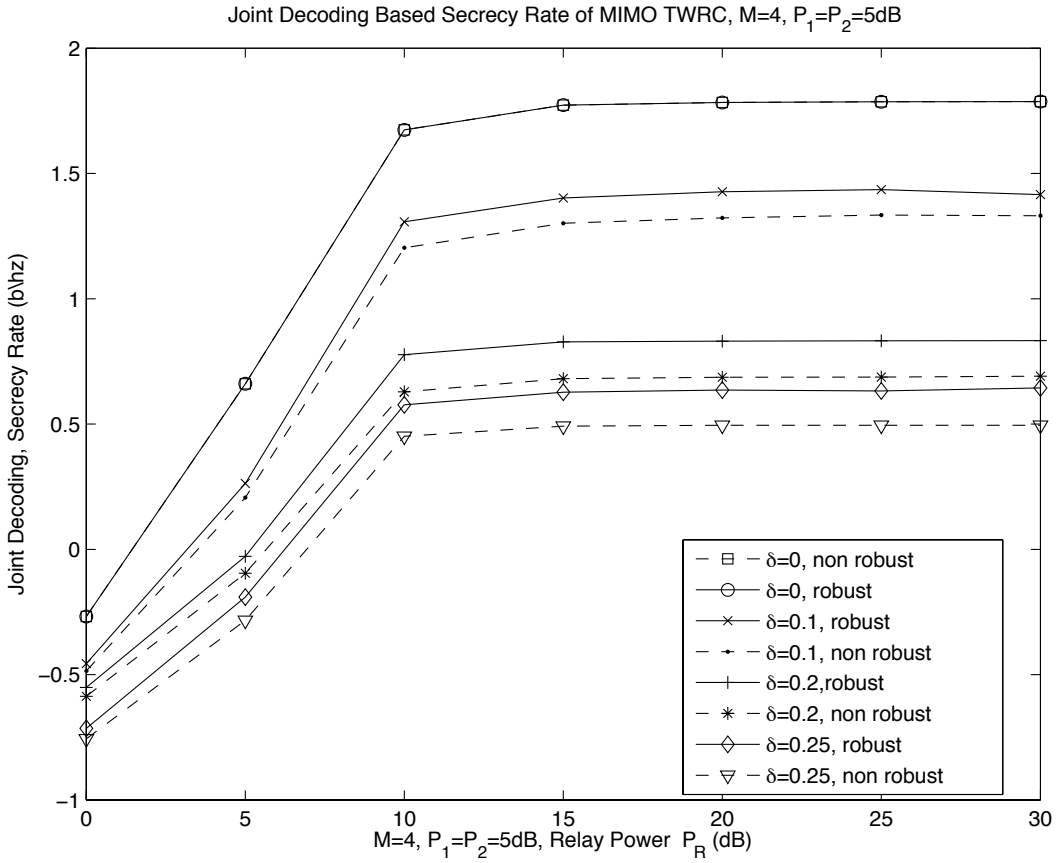

Figure 6.3: Secrecy rate of MIMO TWRC with an untrusted relay, with $M = 4$, $P_1 = P_2 = 5$ dB.

in 6.8 are provided for the required relay transmit power against the secrecy rate requirements. For the results in Fig. 6.7, we investigate the probability that (6.41) is infeasible or the optimal sum rate is non-positive. In the simulations, we considered that $M = 4$, $P_1 = P_2 = 10$ dB, with $\varepsilon = \{0.1, 0.2\}$. It is found that the total relay transmit power required to achieve the positive sum-secrecy rate increases with increase in the CSI error norm bound. The robust beamformer design problem becomes infeasible for the relay power beyond a threshold. From the results, we can observe that this threshold increases with increase in the error norm bound.

On the other hand, in Fig. 6.8, we considered $P_1 = P_2 = 10$ dB and simulated for each (m, P_R) pair for the problem (6.47) with $\alpha_1 r = \alpha_2 r$, for the non-robust perfect CSI ($\varepsilon_{\mathbf{h}}^2 = \varepsilon_{\mathbf{g}}^2 = 0$) network model by using the bisectional searching algorithm. The minimal relay power required is linear with the secrecy rate's threshold value. While with the same secrecy rate constrain threshold value, the more the number of antennas the relay has ,the less power it needed, due to the fact that it has more elements in its beamforming matrix to control the amplitude and angle of the received signals.

6.5 Conclusion

In this chapter, beamforming for AF relaying is studied under secrecy constraints in TWRC. The relay is an unreliable relay with multiple antennas for input and output signal processing. Under this assump-

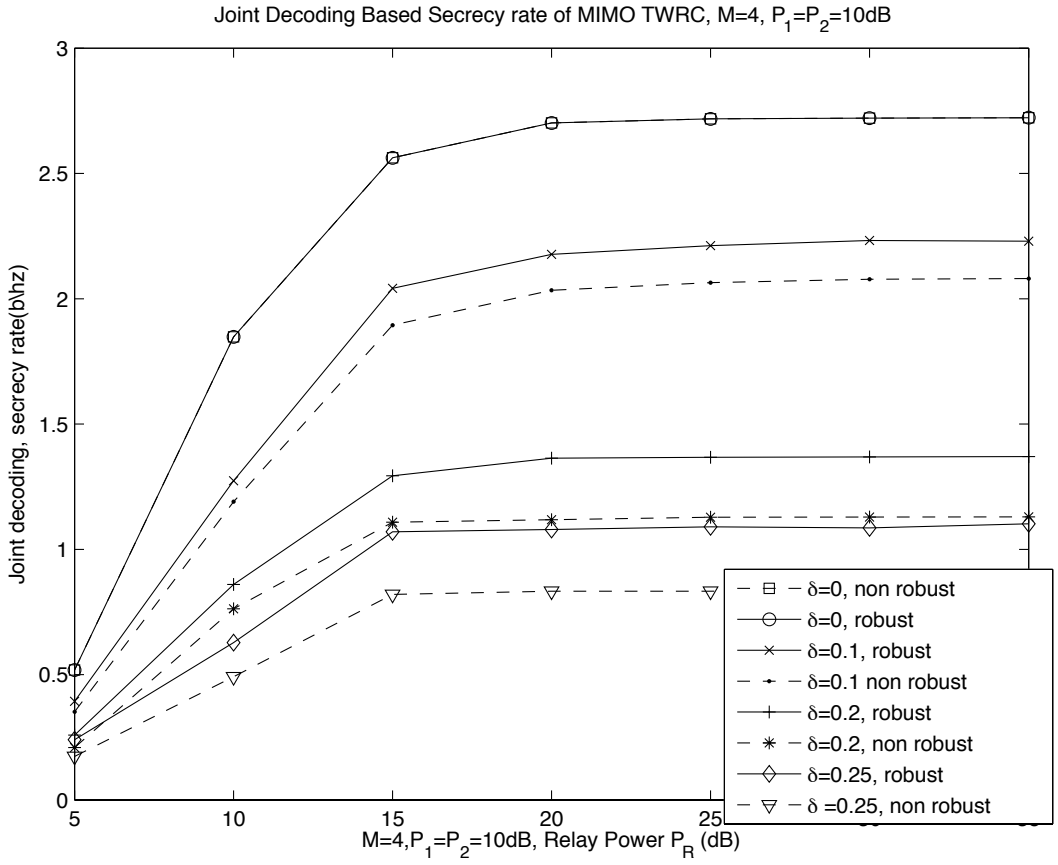

Figure 6.4: Secrecy rate of MIMO TWRC with an untrusted relay, with $M = 4$, $P_1 = P_2 = 10$ dB.

tion, we targeted to optimize the network to achieve a maximal joint-decoding secrecy rate (with the consideration of strict secrecy rate definition of TWRC) with a relay power constraint. We assumed that the allocation of the power to S_1 and S_2 are the same.

For the original non-convex optimization problems, we have presented an SDP formulation, under which optimal beamforming designs that maximize secrecy rates are provided under relay power constraint, with either the cases perfect CSI is available or the case of ellipsoid norm bound ICSI.

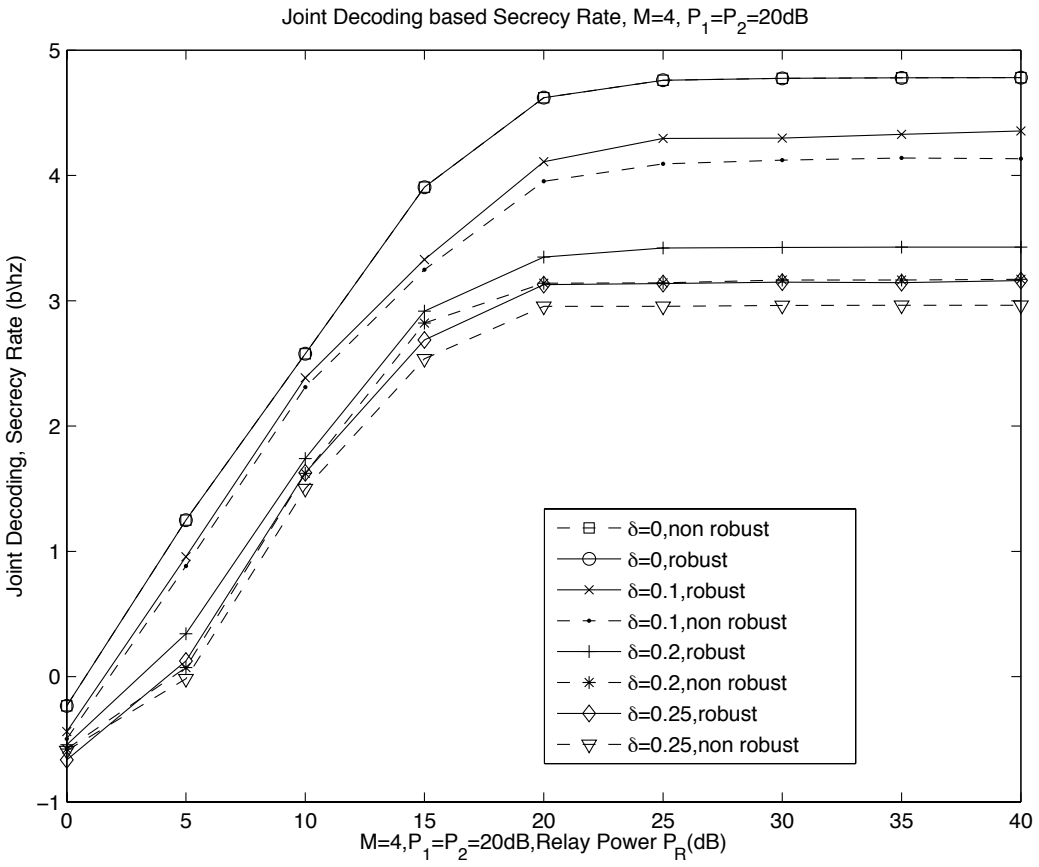

Figure 6.5: Secrecy rate of MIMO TWRC with an untrusted relay, with $M = 4$, $P_1 = P_2 = 20$ dB.

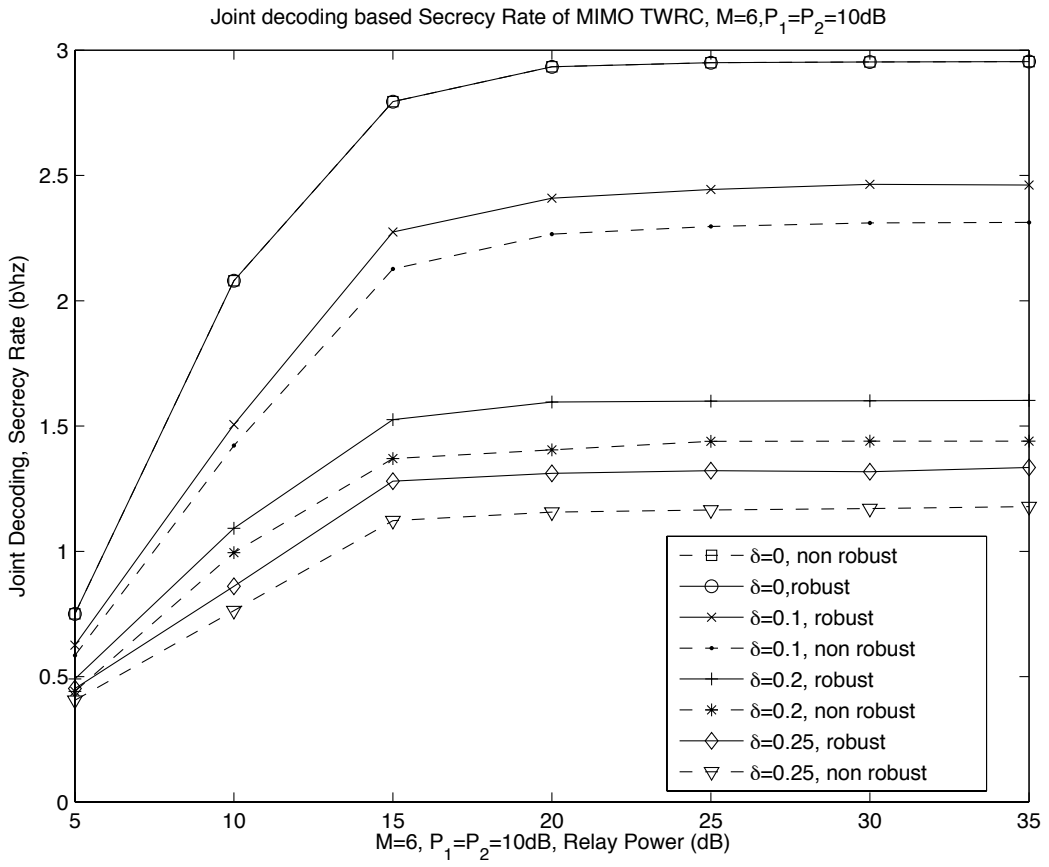

Figure 6.6: Secrecy rate of MIMO TWRC with an untrusted relay, with $M = 6$, $P_1 = P_2 = 10$ dB.

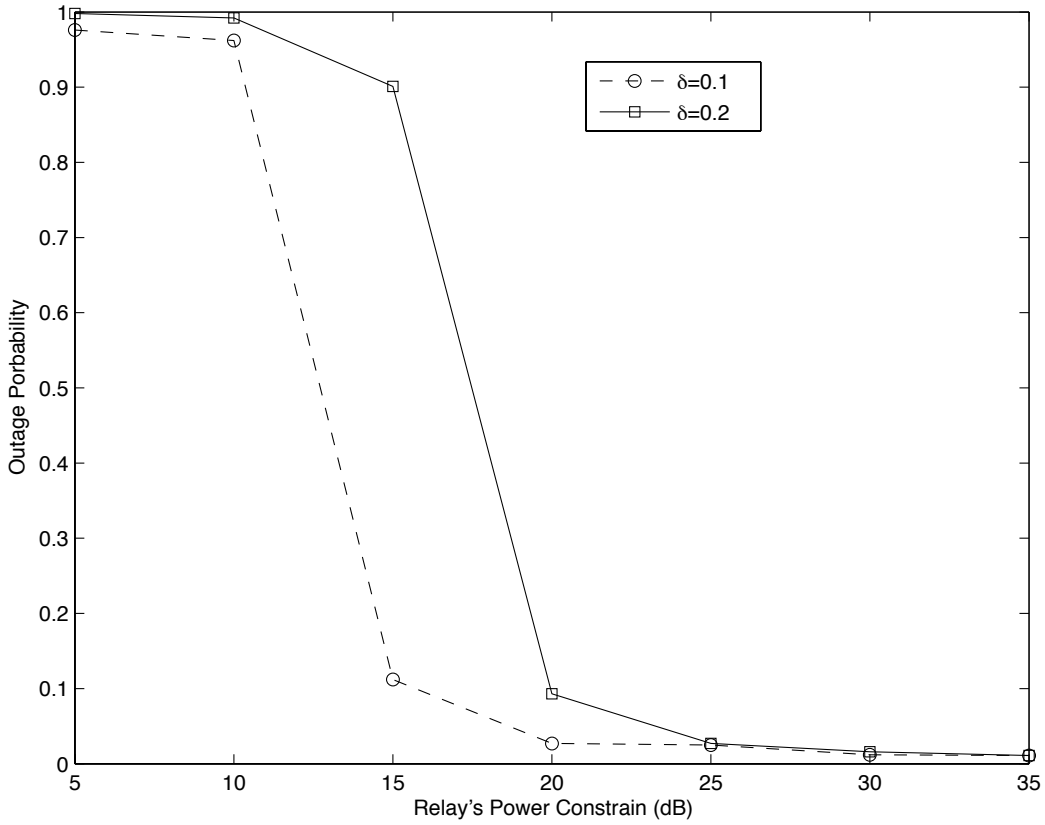

Figure 6.7: Transmit relay power versus the sum secrecy rate.

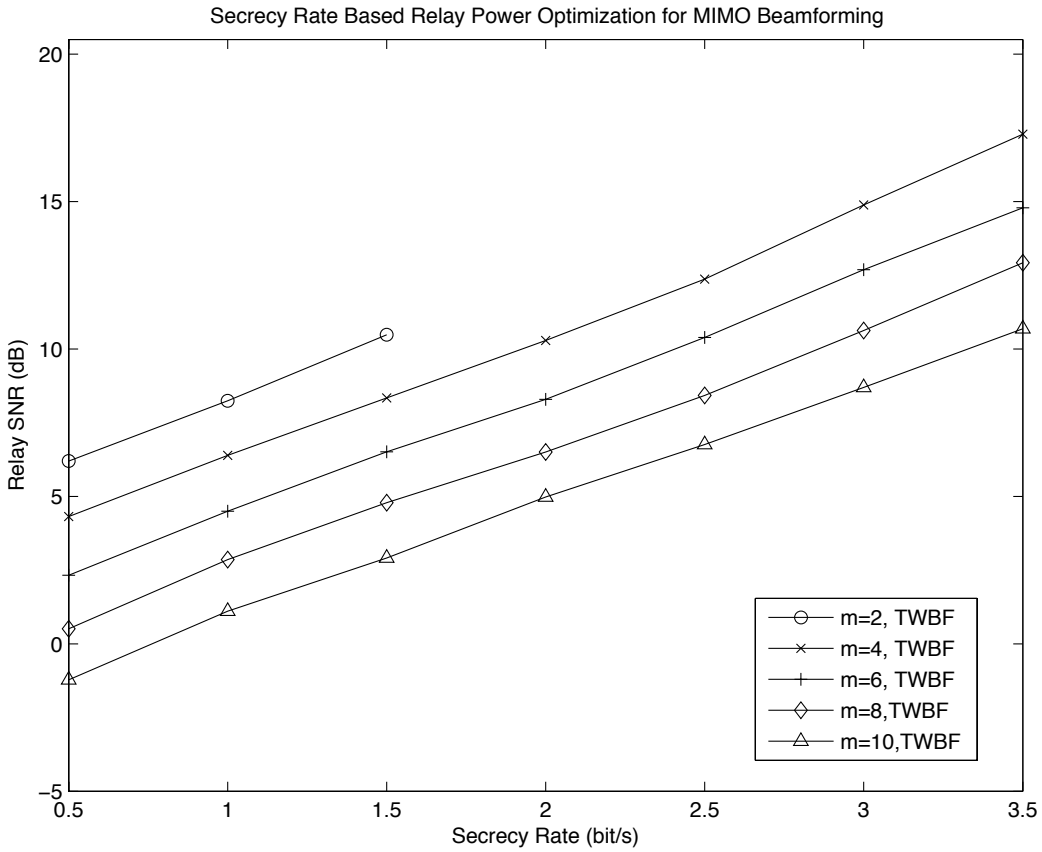

Figure 6.8: Transmit relay SNR versus the secrecy rate requirements.

Chapter 7

Conclusions and Future Works

TWRC is one form of physical layer NC that allows simultaneous transmission of messages. Due to beamforming at the relay, it can largely improve the channel capacity comparing to the OWRC. However, although optimisation of TWRC has been widely studied, for TWRC with a MIMO relay, the secrecy rate based optimisation was not understood, which has set the objective of this thesis. In this thesis, There are two major foci: first the TWRC with perfect CSI and second the TWRC with ICSI, both with consideration of a MIMO relay. Our contributions are summarised in the following section.

7.1 Summary of Contributions

Our contributions are included in Chapters 3–Chapter 6. In Chapter 3, we have investigated the structure of the relay beamforming matrix, if perfect full CSI is available throughout the whole network model. It was proved that the optimal beamforming matrix can be reduced to a rank-2 structure, regardless of the number of antennas at the relay. The outcome is that we are able to provide an SOCP formulation which can obtain the optimal relay beamforming solution at much lower complexity than any existing solutions for the SNR balancing problem of the TWRC.

In Chapter 4, in the presence of an external eavesdropper, we considered the optimisation of the relay beamforming matrix for maximising the secrecy rate assuming perfect CSI. We studied the problem and showed that the optimal relay beamforming matrix has at most rank of 3. Also, we have provided a 2D search method to obtain the optimal beamforming solution and a suboptimal TWZF solution.

In Chapter 5, we further studied the TWRC system but with ICSI where the errors are bounded by ellipsoids. We considered the optimisation of the relay beamforming matrix, with the objective to maximise the joint-secrecy rate. A major effort has been done to rewrite the secrecy constraints with ICSI into LMIs using S-Lemmas. The significant contribution is that it then allows to rewrite the non-convex optimisation problem into an SDP form which after rank relaxation can be solved using standard convex optimisation algorithms. The robust optimal relay beamforming matrix can therefore be obtained for the first time, for the secrecy rate maximisation with ICSI. Chapter 6 repeats the effort but for the case that the relay is the eavesdropper, and the LMIs were derived to obtain the SDP for finding the robust optimal relay beamforming matrix.

7.2 Furture Works

While the solutions derived from convex optimisation are highly attractive, not only because it is optimal, but also that it is computationally efficient. One more steps derived from the robust problems in Chapter 5 and Chapter 6 are closed-form solutions or partial close-form solution for the beamforming matrix of the relay. It would be worth future efforts to investigate whether this is possible in any of the problems studied in this thesis.

From the constrains of the optimisation problems in Chapter 5 and Chapter 6 , a feasible beamformg matrix solution could be derived, and the optimal bemmforming matrix solution could be presented mathematically within this feasible value range. Zero-forcing methods for the optimisation problems in Chapter 5 and Chapter 6 could also been discussed as a low-complexing sub-optimal solution.

Besides, the worst-case approach requires the norms to be bounded, which is usually not satisfied in practice. Also, this approach is too pessimistic since the probability of the worst-case may be extremely low. Hence, statistical approach is a good alternative in certain scenarios.

Taking the optimisation problem in Chapter 4 as an example,

$$\tilde{y}_1 = \sqrt{P_2}\mathbf{h}^T\mathbf{W}(\mathbf{g}+\mathbf{\Delta g})s_2 + \mathbf{h}^T\mathbf{W}\mathbf{v} + \eta_1. \tag{7.1}$$

For S_2, we have

$$\tilde{y}_2 = \sqrt{P_1}\mathbf{g}^T\mathbf{W}(\mathbf{h}+\mathbf{\Delta h})s_1 + \mathbf{g}^T\mathbf{W}\mathbf{v} + \eta_2. \tag{7.2}$$

Similarly for eavesdropper node E, we have

$$\tilde{e} = \sqrt{P_1}\mathbf{e}^T\mathbf{W}(\mathbf{h}+\mathbf{\Delta h})s_1 + \sqrt{P_2}\mathbf{e}^T\mathbf{W}(\mathbf{g}+\mathbf{\Delta g})s_2 + \mathbf{e}^T\mathbf{W}\mathbf{v} + \eta_e. \tag{7.3}$$

By applying the statistical approach, we require the probability of the non-outage for secrecy transmission is greater than the predefined threshold ε by imposing

$$\min_{\mathbf{W}} \; P_1\|\mathbf{W}\mathbf{h}_1\|^2 + P_2\|\mathbf{W}\mathbf{g}_1\|^2 + \mathsf{trace}(\mathbf{W}\mathbf{W}^\dagger)\sigma^2 \tag{7.4a}$$

$$\text{s.t.} \quad \begin{cases} Pr\{\mathsf{R}_{s,1} \geq \Gamma_1\} \geq \varepsilon, \\ Pr\{\mathsf{R}_{s,2} \geq \Gamma_2\} \geq \varepsilon, \end{cases} \tag{7.4b}$$

Statistical approach for robust beamforming for the optimisation problems in chapter 5 and chapter 6 can also be a interesting future research steps for TWRC.

Appendix A

Appendix A

A.1 Appendix I

Given $\mathbf{X}^*$ as the optimal solution of the SDP (A-33), we can apply the algorithm in [77] to decompose $\mathbf{X}^*$ into

$$\mathbf{X}^* = \sum_{i=1}^{r} \mathbf{x}_i \mathbf{x}_i^T \text{ s.t. } \mathbf{x}_i^T(\mathbf{F}_2 - \mathbf{F}_1)\mathbf{x}_i \geq 0 \text{ for } i = 1, \ldots, r, \tag{A-1}$$

where $r = \mathsf{rank}(\mathbf{X}^*)$. Further, define $y_{ij} \triangleq \mathbf{x}_j^T \mathbf{F}_i \mathbf{x}_j$ for $i = 0, 1, 2, 3, 4$ and $j = 1, \ldots, r$. The following linear program

$$\begin{aligned} &\min_{t_1,\ldots,t_r \geq 0} \sum_{j=1}^{r} y_{0j} t_j \\ &\text{s.t.} \begin{cases} \sum_{j=1}^{r} y_{1j} t_j \geq 1, \sum_{j=1}^{r} y_{2j} t_j \geq 1, \\ \sum_{j=1}^{r} y_{3j} t_j = 1 \left(\text{or } \sum_{j=1}^{r} y_{4j} t_j = 1 \right) \end{cases} \end{aligned} \tag{A-2}$$

has the same optimal objective value as (A-33) .

Since $\mathbf{x}_i^T(\mathbf{F}_2 - \mathbf{F}_1)\mathbf{x}_i \geq 0$ for all i, $y_{2j} \geq y_{1j}$ for all j. Therefore, $\sum_{j=1}^{r} y_{1j} t_j \geq 1$ implies $\sum_{j=1}^{r} y_{2j} t_j \geq 1$.

For $r \geq 2$, we can always decrease the number of $\{t_j\}$ to be $r-1$ from the equality constraint and (A-2) becomes

$$\min_{t_1,\ldots,t_r \geq 0} \sum_{j=1}^{r} y_{0j} t_j \text{ s.t. } \sum_{j=1}^{r} y_{1j} t_j \geq 1, \tag{A-3}$$

which was proved in [88] that there is a $t_k > 0$ for $k \leq r$ and $t_j = 0$, for $j \neq k$ and $j \geq 1$ and that there exists an optimal rank-one solution for the SDP problem (A-33). For $r = 1$, the SDP solution is of rank-one and the relaxation is exact.

A.2 Appendix II

In the following sections, we introduce an algorithm to generate a rank-one solution from the optimal solution $\bar{\mathbf{W}}^*$ of (5.97).

As the optimal beamforming matrix $\bar{\mathbf{W}}^*$ is a Hermitian matrix, by Eigen value decomposition,we

have

$$\bar{\mathbf{W}}^* = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^\dagger,$$

After that, we generate a set of vectors $\{\mathbf{w}_k\}$ through the blow process:

$$\mathbf{w}_k = \mathbf{U}\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{v}_k,$$

where $\mathbf{v}_k$ is a circularly symmetric complex Gaussian random variable which has zero mean and an identity covariance matrix. Create $\mathbf{W}^*$ as

$$\mathbf{W}^* = \mathsf{ivec}\mathbf{w}_k$$

Take $\mathbf{W}^*$ into the relay power constraint to check the relay power's feasibility with $\mathbf{W}^*$. Best solution of $\mathbf{w}_k$ is selected through this feasibility check.

We also introduce an algorithm to prove a rank-one solution exists from the optimal solution $\bar{\mathbf{W}}^*$ of (5.97). Consider that the rank of the optimal solution $\bar{\mathbf{W}}^*$ and with $\mathsf{rank}(\bar{\mathbf{W}}^*) = r$. First, we decompose $\bar{\mathbf{W}}^*$ as

$$\begin{aligned} \text{find } \bar{\mathbf{W}}^* &= \sum_{i=1}^{r} \mathbf{w}_i \mathbf{w}_i^T \\ \text{s.t. } &\begin{cases} \mathbf{w}_i^T \mathbf{F}_1 \mathbf{w}_i \geq 0, \\ \mathbf{w}_i^T \mathbf{F}_2 \mathbf{w}_i \geq 0, \\ \mathbf{w}_i^T \mathbf{F}_e \mathbf{w}_i \geq 0, \\ \mathbf{w}_i^T \mathbf{F}_R \mathbf{w}_i \geq P_R, \end{cases} \end{aligned} \tag{A-4}$$

where $r = \mathsf{rank}(\bar{\mathbf{W}}^*)$. Further, define $y_{ij} \triangleq \mathbf{w}_j^T \mathbf{F}_i \mathbf{w}_j$ for $i = 0, 1, 2, 3, 4$ and $j = 1, \ldots, r$. The following linear program

$$\begin{aligned} &\min_{t_1,\ldots,t_r \geq 0} \sum_{j=1}^{r} y_{0j} t_j \\ &\text{s.t. } \begin{cases} \sum_{j=1}^{r} y_{1j} t_j \geq 1, \sum_{j=1}^{r} y_{2j} t_j \geq 1, \\ \sum_{j=1}^{r} y_{3j} t_j = 1 \left(\text{or } \sum_{j=1}^{r} y_{4j} t_j = 1 \right) \end{cases} \end{aligned} \tag{A-5}$$

has the same optimal objective value as (5.97). Since $\mathbf{w}_i^T(\mathbf{F}_2 - \mathbf{F}_1)\mathbf{w}_i \geq 0$ for all i, $y_{2j} \geq y_{1j}$ for all j. Therefore, $\sum_{j=1}^{r} y_{1j} t_j \geq 1$ implies $\sum_{j=1}^{r} y_{2j} t_j \geq 1$. For $r \geq 2$, we can always decrease the number of $\{t_j\}$ to be $r-1$ from the equality constraint and (A-5) becomes

$$\min_{t_1,\ldots,t_r \geq 0} \sum_{j=1}^{r} y_{0j} t_j \text{ s.t. } \sum_{j=1}^{r} y_{1j} t_j \geq 1. \tag{A-6}$$

A.3 Appendix III

To do so, we set $\mathbf{A} = [\mathbf{h}\ \mathbf{g}\ \mathbf{e}] \in \mathcal{C}^{M\times 3}$. Then we write the SVD of $\mathbf{A}$ as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\dagger}, \tag{A-7}$$

where $\mathbf{U} = [\mathbf{U}^{\parallel}\ \mathbf{U}^{\perp}] \in \mathcal{C}^{M\times M}$ is a unitary matrix with $\mathbf{U}^{\parallel} \in \mathcal{C}^{M\times 3}$ and $\mathbf{U}^{\perp} \in \mathcal{C}^{M\times(M-3)}$, $\mathbf{V} \in \mathcal{C}^{3\times 3}$ is another unitary matrix and

$$\mathbf{\Sigma} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix} \in \mathcal{C}^{M\times 3}, \tag{A-8}$$

in which $\lambda_1 \geq \lambda_2 \geq 0$ are the singular values of $\mathbf{A}$. As a consequence, $\mathbf{W}$ can be expressed as

$$\begin{aligned} \mathbf{W} =& [(\mathbf{U}^{\parallel})^*\ (\mathbf{U}^{\perp})^*] \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{D} & \mathbf{E} \end{bmatrix} [\mathbf{U}^{\parallel}\ \mathbf{U}^{\perp}]^{\dagger} \\ =& (\mathbf{U}^{\parallel})^*\mathbf{B}(\mathbf{U}^{\parallel})^{\dagger} + (\mathbf{U}^{\parallel})^*\mathbf{C}(\mathbf{U}^{\perp})^{\dagger} + (\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\parallel})^{\dagger} + (\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^{\dagger}. \end{aligned} \tag{A-9}$$

From (A-9), we can get

$$\begin{aligned} & |\mathbf{h}^T\mathbf{W}\mathbf{g}|^2 \\ =& |\mathbf{h}^T(\mathbf{U}^{\parallel})^*\mathbf{B}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{g} + \mathbf{h}^T(\mathbf{U}^{\parallel})^*\mathbf{C}(\mathbf{U}^{\perp})^{\dagger}\mathbf{g} + \mathbf{h}^T(\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{g} + \mathbf{h}^T(\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^{\dagger}\mathbf{g}|^2 \\ =& |\mathbf{h}^T(\mathbf{U}^{\parallel})^*\mathbf{B}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{g}|^2, \end{aligned} \tag{A-10}$$

$$\begin{aligned} & |\mathbf{g}^T\mathbf{W}\mathbf{h}|^2 \\ =& |\mathbf{g}^T(\mathbf{U}^{\parallel})^*\mathbf{B}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{h} + \mathbf{g}^T(\mathbf{U}^{\parallel})^*\mathbf{C}(\mathbf{U}^{\perp})^{\dagger}\mathbf{h} + \mathbf{g}^T(\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{h} + \mathbf{g}^T(\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^{\dagger}\mathbf{h}|^2 \\ =& |\mathbf{g}^T(\mathbf{U}^{\parallel})^*\mathbf{B}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{h}|^2 \end{aligned} \tag{A-11}$$

$$\begin{aligned} & |\mathbf{e}^T\mathbf{W}\mathbf{h}|^2 \\ =& |\mathbf{e}^T(\mathbf{U}^{\parallel})^*\mathbf{B}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{h} + \mathbf{e}^T(\mathbf{U}^{\parallel})^*\mathbf{C}(\mathbf{U}^{\perp})^{\dagger}\mathbf{h} + \mathbf{e}^T(\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{h} + \mathbf{e}^T(\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^{\dagger}\mathbf{h}|^2 \\ =& |\mathbf{e}^T(\mathbf{U}^{\parallel})^*\mathbf{B}(\mathbf{U}^{\parallel})^{\dagger}\mathbf{h}|^2 \end{aligned} \tag{A-12}$$

$$
\begin{aligned}
&|\mathbf{e}^T\mathbf{W}\mathbf{g}|^2 \\
=&|\mathbf{e}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger\mathbf{g}+\mathbf{e}^T(\mathbf{U}^{\|})^*\mathbf{C}(\mathbf{U}^{\perp})^\dagger\mathbf{g}+\mathbf{e}^T(\mathbf{U}^{\perp})^*\mathbf{D}(\mathbf{U}^{\|})^\dagger\mathbf{g}+\mathbf{e}^T(\mathbf{U}^{\perp})^*\mathbf{E}(\mathbf{U}^{\perp})^\dagger\mathbf{g}|^2 \\
=&|\mathbf{e}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger\mathbf{g}|^2. \qquad \text{(A-13)}
\end{aligned}
$$

Also, we have

$$
\begin{aligned}
\|\mathbf{W}^\dagger\mathbf{h}^*\|^2 &= \|\mathbf{B}^\dagger(\mathbf{U}^{\|})^T\mathbf{h}^*\|^2+\|\mathbf{C}^\dagger(\mathbf{U}^{\|})^T\mathbf{h}^*\|^2+\|\mathbf{D}^\dagger\mathbf{U}^{\perp}\mathbf{h}^*\|^2+\|\mathbf{D}^\dagger\mathbf{U}^{\perp}\mathbf{h}^*\|^2 \\
&= \|\mathbf{B}^\dagger\mathbf{U}^T\mathbf{h}^*\|^2+\|\mathbf{C}^\dagger\mathbf{U}^T\mathbf{h}^*\|^2 \qquad \text{(A-14)}
\end{aligned}
$$

$$
\begin{aligned}
\|\mathbf{W}^\dagger\mathbf{g}^*\|^2 &= \|\mathbf{B}^\dagger(\mathbf{U}^{\|})^T\mathbf{g}^*\|^2+\|\mathbf{C}^\dagger(\mathbf{U}^{\|})^T\mathbf{g}^*\|^2+\|\mathbf{D}^\dagger\mathbf{U}^{\perp}\mathbf{g}^*\|^2+\|\mathbf{D}^\dagger\mathbf{U}^{\perp}\mathbf{g}^*\|^2 \\
&= \|\mathbf{B}^\dagger(\mathbf{U}^{\|})^T\mathbf{g}^*\|^2+\|\mathbf{C}^\dagger(\mathbf{U}^{\|})^T\mathbf{g}^*\|^2 \qquad \text{(A-15)}
\end{aligned}
$$

$$
\begin{aligned}
\|\mathbf{W}^\dagger\mathbf{e}^*\|^2 &= \|\mathbf{B}^\dagger(\mathbf{U}^{\|})^T\mathbf{e}^*\|^2+\|\mathbf{C}^\dagger(\mathbf{U}^{\|})^T\mathbf{g}^*\|^2+\|\mathbf{D}^\dagger\mathbf{U}^{\perp}\mathbf{e}^*\|^2+\|\mathbf{D}^\dagger\mathbf{U}^{\perp}\mathbf{e}^*\|^2 \\
&= \|\mathbf{B}^\dagger(\mathbf{U}^{\|})^T\mathbf{e}^*\|^2+\|\mathbf{C}^\dagger(\mathbf{U}^{\|})^T\mathbf{h}\mathbf{e}^*\|^2, \qquad \text{(A-16)}
\end{aligned}
$$

where $(\cdot)^*$ represents the conjugate operation, and $\mathbf{B}\in\mathcal{C}^{3\times3}$, $\mathbf{C}\in\mathcal{C}^{3\times(M-3)}$, $\mathbf{D}\in\mathcal{C}^{(M-3)\times3}$ and $\mathbf{E}\in\mathcal{C}^{(M-3)\times(M-3)}$ are matrices of appropriate sizes. Using this structure onto the constraints in (4.11), it can be easily observed that they are not related to $\mathbf{D}$ and $\mathbf{E}$ and further, for minimising the relaying power, $\mathbf{D}$ and $\mathbf{E}$ should all be set to zeros. Also, consider the influence of $\mathbf{C}$ for R_{s1}, R_{s2} and P_R which is proved in detail in Appendix IV, $\mathbf{C}$ should not be 0. Taking R_{s1} as example,

$$
\begin{aligned}
R_{s1} = &\frac{1}{2}\log(1+\underbrace{\frac{P_2|\mathbf{h}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger\mathbf{g}|^2}{\sigma^2(\|\mathbf{B}^\dagger\mathbf{U}^T\mathbf{h}^*\|^2+\|\mathbf{C}^\dagger\mathbf{U}^T\mathbf{h}^*\|^2+1)}}_{A_1}) \\
&-\frac{1}{2}\log(1+\underbrace{\frac{P_1(|\mathbf{e}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger\mathbf{g}|^2+|\mathbf{g}_e|^2)}{P_1(|\mathbf{e}^T(\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger\mathbf{h}|^2+|\mathbf{h}_e|^2)+\sigma^2(\|\mathbf{B}^\dagger\mathbf{U}^T\mathbf{e}^*\|^2+\|\mathbf{C}^\dagger\mathbf{U}^T\mathbf{e}^*\|^2+2)}}_{A_2}).
\end{aligned} \qquad \text{(A-17)}
$$

Then set

$$\mathbf{Q} = (\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger, \qquad \text{(A-18)}$$

$$\mathbf{q} = \mathsf{vec}(\mathbf{Q}), \qquad \text{(A-19)}$$

$$\mathbf{K} = (\mathbf{U}^{\|})^*\mathbf{C}(\mathbf{U}^{\|})^\dagger, \qquad \text{(A-20)}$$

$$\mathbf{k} = \mathsf{vec}(\mathbf{K}). \qquad \text{(A-21)}$$

Taking the above definition into A_1 of (A-17), we get

$$A_1 = \frac{P_2|\mathbf{f}_1^T q|^2}{\sigma^2(\|\mathbf{Hq}\|^2 + \|\mathbf{Hk}\|^2 + 1)}, \tag{A-22}$$

$$A_2 = \frac{P_1(|\mathbf{f}_4^T q|^2 + g_e^2)}{P_1(|\mathbf{f}_3^T q|^2 + h_e^2) + \sigma^2(\|\mathbf{Eq}\|^2 + \|\mathbf{Ek}\|^2 + 2)}. \tag{A-23}$$

Now, letting

$$a^2 = \|\mathbf{Hk}\|^2, \text{ for } a \geq 0, \tag{A-24}$$

$$b^2 = \|\mathbf{Ek}\|^2, \tag{A-25}$$

$$\mathbf{q}_1 = [\mathbf{q}, 1], \tag{A-26}$$

$$\mathbf{K}_1 \triangleq \begin{bmatrix} P_2\mathbf{f}_1 * \mathbf{f}_1^T & 0 \\ 0 & 0 \end{bmatrix}, \tag{A-27}$$

$$\mathbf{K}_2 \triangleq \sigma^2 \begin{bmatrix} \mathbf{H}^\dagger\mathbf{H} & 0 \\ 0 & a^2 + 1 \end{bmatrix}, \tag{A-28}$$

$$\mathbf{K}_3 \triangleq \begin{bmatrix} P_1\mathbf{f}_4 * \mathbf{f}_4^T & 0 \\ 0 & g_e^2 \end{bmatrix}, \tag{A-29}$$

$$\mathbf{K}_4 \triangleq \begin{bmatrix} P_1\mathbf{f}_3 * \mathbf{f}_3^T + \sigma^2\mathbf{E}^\dagger\mathbf{E} & 0 \\ 0 & h_e^2 P_1 + \sigma^2 b^2 + 2\sigma^2 \end{bmatrix}, \tag{A-30}$$

we get

$$A_1 = \frac{\mathbf{q}_1^T\mathbf{K}_1\mathbf{q}_1}{\mathbf{q}_1^T\mathbf{K}_2\mathbf{q}_1}, \tag{A-31}$$

$$A_2 = \frac{\mathbf{q}_1^T\mathbf{K}_3\mathbf{q}_1}{\mathbf{q}_1^T\mathbf{K}_4\mathbf{q}_1}. \tag{A-32}$$

Let $\mathbf{L}$ be the Cholesky decomposition of $\mathbf{H}^\dagger\mathbf{H}$, $\mathbf{L}_1$ be the Cholesky decomposition of $\mathbf{K}_2$, $\mathbf{L}_3$ be the Cholesky decomposition of matrix $P_1\mathbf{f}_3 * \mathbf{f}_3^T + \sigma^2\mathbf{E}^\dagger\mathbf{E}$, and $\mathbf{L}_4$ be the Cholesky decomposition of $\mathbf{K}_4$. Then we have

$$\mathbf{L}_1 \triangleq \sigma \begin{bmatrix} \mathbf{L} & 0 \\ 0 & \sqrt{a^2 + 1} \end{bmatrix} \tag{A-33}$$

and

$$\mathbf{L}_1^{-1} \triangleq \sigma \begin{bmatrix} \mathbf{L}^{-1} & 0 \\ 0 & \frac{1}{\sqrt{a^2+1}} \end{bmatrix}. \tag{A-34}$$

Also, define

$$\mathbf{L}_2 \triangleq \mathbf{L}_1^{*-1}\mathbf{K}_1\mathbf{L}_1^{-1} \triangleq \begin{bmatrix} P_2\sigma\mathbf{H}^\dagger\mathbf{H}\mathbf{f}_1 * \mathbf{f}_1^T & 0 \\ 0 & 0 \end{bmatrix} \tag{A-35}$$

$\mathbf{y} = \mathbf{L}_1\mathbf{q}_1$, and

$$\mathbf{L}_3 \triangleq \sigma \begin{bmatrix} \mathbf{L}_3 & 0 \\ 0 & \sqrt{h_e^2 P_1 + \sigma^2 b^2 + 2\sigma^2} \end{bmatrix} \tag{A-36}$$

and

$$\mathbf{L}_4^{-1} \triangleq \sigma \begin{bmatrix} \mathbf{L}_3^{-1} & 0 \\ 0 & \frac{1}{\sqrt{h_e^2 P_1 + \sigma^2 b^2 + 2\sigma^2}} \end{bmatrix}. \tag{A-37}$$

Furthermore, define

$$\mathbf{L}_2 \triangleq \mathbf{L}_1^{*-1}\mathbf{K}_1\mathbf{L}_1^{-1} \triangleq \begin{bmatrix} P_2\sigma\mathbf{H}^\dagger\mathbf{H}\mathbf{f}_1 * \mathbf{f}_1^T & 0 \\ 0 & 0 \end{bmatrix} \tag{A-38}$$

and $\mathbf{z} = \mathbf{L}_4\mathbf{q}_1$. As a result, we get

$$A_1 = \frac{\mathbf{y}^T\mathbf{L}_2\mathbf{y}}{\mathbf{y}^T\mathbf{y}}, \tag{A-39}$$

$$A_2 = \frac{\mathbf{z}^T\mathbf{L}_4\mathbf{z}}{\mathbf{z}^T\mathbf{z}}. \tag{A-40}$$

Since $\mathbf{L}_2$ is not related to a, and $\mathbf{L}_4$ is related to b, we are unable to set $\mathbf{C} = 0$. Hence,

$$\mathbf{W} = (\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\|})^\dagger + (\mathbf{U}^{\|})^*\mathbf{B}(\mathbf{U}^{\perp})^\dagger. \tag{A-41}$$

Thus the rank of optimal matrix is generally $\mathsf{rank}(\mathbf{W}) = 3$.

A.4 Appendix IV

From the zero-forcing constraint for the eavesdropper's signal's in (4.22), i.e., $\mathbf{e}^T\mathbf{W} = 0$, and as $\mathbf{W} \in \mathcal{C}^{M\times M}$, we can conclude that $\mathsf{rank}(\mathbf{W}) = 1$ and every vector of $\mathbf{W}$ is perpendicular to $\mathbf{e}_1$. Define $\mathbf{e}_1^\perp = [e_1\ e_2]^T$, such that $\mathbf{e}_1^T\mathbf{e}_1^\perp = 0$. Without loss of generosity, $\|\mathbf{e}_1^\perp\|^2 = 1$. As a result, we can rewrite $\mathbf{W}$ as:

$$\mathbf{W} = \begin{bmatrix} l\mathbf{e}_1^\perp & k\mathbf{e}_1^\perp \end{bmatrix} \tag{A-42}$$

where l, k are complex coefficient numbers.

Now, define $\mathbf{h}_1 \triangleq (\mathbf{U}^{\|})^\dagger\mathbf{h} \in \mathcal{C}^{3\times 1}$, $\mathbf{g}_1 \triangleq (\mathbf{U}^{\|})^\dagger\mathbf{g} \in \mathcal{C}^{3\times 1}$ and $\mathbf{e}_1 \triangleq (\mathbf{U}^{\|})^\dagger\mathbf{e} \in \mathcal{C}^{3\times 1}$. Take (A-42)

and consider:

$$
\begin{aligned}
a(\mathbf{W}) &= |\mathbf{h}_1^T\mathbf{W}\mathbf{g}_1|^2 = \left|\mathbf{h}_1^T \begin{bmatrix} l\mathbf{e}_1^\perp & k\mathbf{e}_1^\perp \end{bmatrix} \mathbf{g}_1\right|^2 \\
&= \left|\begin{bmatrix} l\mathbf{h}_1^T\mathbf{e}_1^\perp & k\mathbf{h}_1^T\mathbf{e}_1^\perp \end{bmatrix} \mathbf{g}_1\right|^2 \\
&= (lg_{11}\mathbf{h}_1^T\mathbf{e}_1^\perp + kg_{21}\mathbf{h}_1^T\mathbf{e}_1^\perp)^2 \\
&= (\mathbf{h}_1^T\mathbf{e}_1^\perp)^2(lg_{11} + kg_{21})^2 \qquad \text{(A-43)} \\
b(\mathbf{W}) &= |\mathbf{g}_1^T\mathbf{W}\mathbf{h}_1|^2 = \left|\mathbf{g}_1^T \begin{bmatrix} l\mathbf{e}_1^\perp & k\mathbf{e}_1^\perp \end{bmatrix} \mathbf{h}_1\right|^2 \\
&= \left|\begin{bmatrix} l\mathbf{g}_1^T\mathbf{e}_1^\perp & k\mathbf{g}_1^T\mathbf{e}_1^\perp \end{bmatrix} \mathbf{h}_1\right|^2 \\
&= (lh_{11}\mathbf{g}_1^T\mathbf{e}_1^\perp + kh_{21}\mathbf{g}_1^T\mathbf{e}_1^\perp)^2 \\
&= (\mathbf{g}_1^T\mathbf{e}_1^\perp)^2(lh_{11} + kh_{21})^2 \qquad \text{(A-44)} \\
c(\mathbf{W}) &= \sigma^2\left(\|\mathbf{h}_1^T\mathbf{W}\|^2 + 1\right) = \sigma^2\left(\left\|\begin{bmatrix} l\mathbf{h}_1^T\mathbf{e}_1^\perp & k\mathbf{h}_1^T\mathbf{e}_1^\perp \end{bmatrix}\right\|^2 + 1\right) \\
&= \sigma^2\left((l\mathbf{h}_1^T\mathbf{e}_1^\perp)^2 + (k\mathbf{h}_1^T\mathbf{e}_1^\perp)^2 + 1\right) \qquad \text{(A-45)} \\
d(\mathbf{W}) &= \sigma^2\left(\|\mathbf{g}_1^T\mathbf{W}\|^2 + 1\right) = \sigma^2\left(\left\|\begin{bmatrix} l\mathbf{g}_1^T\mathbf{e}_1^\perp & k\mathbf{g}_1^T\mathbf{e}_1^\perp \end{bmatrix}\right\|^2 + 1\right) \\
&= \sigma^2\left((l\mathbf{g}_1^T\mathbf{e}_1^\perp)^2 + (k\mathbf{g}_1^T\mathbf{e}_1^\perp)^2 + 1\right) \qquad \text{(A-46)} \\
x(\mathbf{W}) &= \|\mathbf{W}\mathbf{h}_1\|^2 = \| \begin{bmatrix} l\mathbf{e}_1^\perp & k\mathbf{e}_1^\perp \end{bmatrix} \mathbf{h}_1\|^2 \\
&= \left\| \begin{bmatrix} le_1 & ke_1 \\ le_2 & ke_2 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{21} \end{bmatrix} \right\|^2 \\
&= (le_1h_{11} + ke_1h_{21})^2 + (le_2h_{11} + ke_2h_{21})^2 \\
&= (e_1^2 + e_2^2)(lh_{11} + kh_{21})^2 \qquad \text{(A-47)} \\
y(\mathbf{W}) &= \|\mathbf{W}\mathbf{g}_1\|^2 = \| \begin{bmatrix} l\mathbf{e}_1^\perp & k\mathbf{e}_1^\perp \end{bmatrix} \mathbf{g}_1\|^2 \\
&= \left\| \begin{bmatrix} le_1 & ke_1 \\ le_2 & ke_2 \end{bmatrix} \begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix} \right\|^2 \\
&= (le_1g_{11} + ke_1g_{21})^2 + (le_2g_{11} + ke_2g_{21})^2 \\
&= (e_1^2 + e_2^2)(lg_{11} + kg_{21})^2. \qquad \text{(A-48)}
\end{aligned}
$$

Also, we can get $\frac{x(\mathbf{W})a(\mathbf{W})d(\mathbf{W})}{y(\mathbf{W}^o)b(\mathbf{W})c(\mathbf{W})}$ as

$$
\begin{aligned}
\frac{x(\mathbf{W})a(\mathbf{W})d(\mathbf{W})}{y(\mathbf{W})b(\mathbf{W})c(\mathbf{W})} &= \frac{(e_1^2 + e_2^2)(lh_{11} + kh_{21})^2(\mathbf{h}_1^T\mathbf{e}_1^\perp)^2(lg_{11} + kg_{21})^2\left((l\mathbf{g}_1^T\mathbf{e}_1^\perp)^2 + (k\mathbf{g}_1^T\mathbf{e}_1^\perp)^2 + 1\right)}{(e_1^2 + e_2^2)(lg_{11} + kg_{21})^2(\mathbf{g}_1^T\mathbf{e}_1^\perp)^2(lh_{11} + kh_{21})^2\left((l\mathbf{h}_1^T\mathbf{e}_1^\perp)^2 + (k\mathbf{h}_1^T\mathbf{e}_1^\perp)^2 + 1\right)} \\
&= \frac{(\mathbf{h}_1^T\mathbf{e}_1^\perp)^2\left((\mathbf{g}_1^T\mathbf{e}_1^\perp)^2(l^2 + k^2) + 1\right)}{(\mathbf{g}_1^T\mathbf{e}_1^\perp)^2\left((\mathbf{h}_1^T\mathbf{e}_1^\perp)^2(l^2 + k^2) + 1\right)} \\
&\equiv f_1
\end{aligned}
\qquad \text{(A-49)}
$$

and

$$\begin{aligned}\frac{a(\mathbf{W})}{y(\mathbf{W})c(\mathbf{W})} &= \frac{(\mathbf{h}_1^T\mathbf{e}_1^{\perp})^2(lg_{11}+kg_{21})^2}{\sigma^2(e_1^2+e_2^2)(lg_{11}+kg_{21})^2\left((l\mathbf{h}_1^T\mathbf{e}_1^{\perp})^2+(k\mathbf{h}_1^T\mathbf{e}_1^{\perp})^2+1\right)} \\ &= \frac{(\mathbf{h}_1^T\mathbf{e}_1^{\perp})^2}{\sigma^2\left((\mathbf{h}_1^T\mathbf{e}_1^{\perp})^2(l^2+k^2)+1\right)} \\ &\equiv f_2.\end{aligned} \tag{A-50}$$

As $l^2+k^2=\|\mathbf{W}\|$, without loss of generality, we normalised $\mathbf{W}$ as $l^2+k^2=1$ to get

$$R_s^{'o} = \frac{1}{4f_1}((P_T-\sigma^2)f_2+f_1+1)^2. \tag{A-51}$$

A.5 Appendix V

Af first, define

$$\mathbf{x} = \mathsf{vec}(\mathbf{B}) \tag{A-52}$$

$$\mathbf{f}_1 \triangleq \mathsf{vec}(\mathbf{h}_1\mathbf{g}_1^T) \tag{A-53}$$

$$\mathbf{f}_2 \triangleq \mathsf{vec}(\mathbf{g}_1\mathbf{h}_1^T) \tag{A-54}$$

$$\mathbf{U}_1 \triangleq P_1\mathbf{h}_1\mathbf{h}_1^{\dagger}+P_2\mathbf{g}_1\mathbf{g}_1^{\dagger}+\sigma^2\mathbf{I} \tag{A-55}$$

$$\mathbf{U}_2 \triangleq [\mathsf{diag}(\mathbf{U}_1^T,\mathbf{U}_1^T)]^{\frac{1}{2}} \tag{A-56}$$

$$\mathbf{H} \triangleq \begin{bmatrix} [\mathbf{h}_1]_1 & 0 & [\mathbf{h}_1]_2 & 0 \\ 0 & [\mathbf{h}_1]_1 & 0 & [\mathbf{h}_1]_2 \end{bmatrix} \tag{A-57}$$

$$\mathbf{G} \triangleq \begin{bmatrix} [\mathbf{g}_1]_1 & 0 & [\mathbf{g}_1]_2 & 0 \\ 0 & [\mathbf{g}_1]_1 & 0 & [\mathbf{g}_1]_2 \end{bmatrix} \tag{A-58}$$

where the notation "$[\mathbf{a}]_n$" returns the nth entry of $\mathbf{a}$ (a similar notation is also used for denoting the entry of a matrix). Thus, (6.43) becomes

$$\mathbb{P} \mapsto \begin{cases} \max\limits_{\mathbf{x}} \ \log\left\{1+\dfrac{P_2|\mathbf{f}_1^T\mathbf{x}|^2}{\sigma^2\left(\|\mathbf{H}\mathbf{x}\|^2+1\right)}\right\}+\log\left\{1+\dfrac{P_1|\mathbf{f}_2^T\mathbf{x}|^2}{\sigma^2\left(\|\mathbf{G}\mathbf{x}\|^2+1\right)}\right\} \\ \text{s.t. } \|\mathbf{U}_2\mathbf{x}\|^2 \le P_R. \end{cases} \tag{A-59}$$

By using the decomposition methods described at Appendix IV, we could rewrite $1+\frac{P_2|\mathbf{f}_1^T\mathbf{x}|^2}{\sigma^2(\|\mathbf{H}\mathbf{x}\|^2+1)}$ and $1+\frac{P_1|\mathbf{f}_2^T\mathbf{x}|^2}{\sigma^2(\|\mathbf{G}\mathbf{x}\|^2+1)}$ into the format as

$$f(\mathbf{x}_1) = 1+\frac{P_2|\mathbf{f}_1^T\mathbf{x}|^2}{\sigma^2\left(\|\mathbf{H}\mathbf{x}\|^2+1\right)} = \frac{\mathbf{x}^{\dagger}\mathbf{B}_2\mathbf{x}}{\mathbf{x}^{\dagger}\mathbf{B}_1\mathbf{x}+\sigma^2} = \frac{\mathbf{x}_1^{\dagger}\mathbf{E}_1\mathbf{x}_1}{\mathbf{x}_1^{\dagger}\mathbf{x}_1}, \tag{A-60}$$

$$f(\mathbf{x}_2) = 1+\frac{P_1|\mathbf{f}_2^T\mathbf{x}|^2}{\sigma^2\left(\|\mathbf{G}\mathbf{x}\|^2+1\right)} = \frac{\mathbf{x}^{\dagger}\mathbf{B}_4\mathbf{x}}{\mathbf{x}^{\dagger}\mathbf{B}_3\mathbf{x}+\sigma^2} = \frac{\mathbf{x}_2^{\dagger}\mathbf{E}_2\mathbf{x}_2}{\mathbf{x}_2^{\dagger}\mathbf{x}_2}, \tag{A-61}$$

where $\mathbf{x}_1$, $\mathbf{x}_2$,$\mathbf{B}_i, i = 1, \ldots, 4$, $\mathbf{E}_1$and $\mathbf{E}_1$ are defined following Appendix IV's generalised Raleigh quotient transformation. Suppose that after the transformation, we have

$$\mathbf{x}_1 = \mathbf{S}_1^{\frac{1}{2}}\mathbf{x}, \tag{A-62}$$

$$\mathbf{x}_2 = \mathbf{S}_2^{\frac{1}{2}}\mathbf{x}. \tag{A-63}$$

Then we define $\Xi = \mathbf{S}_1^{-1}\mathbf{B}_2 + \mathbf{S}_2^{-1}\mathbf{B}_4$, and based on [44], a suboptimal solution of (A-59) can be derived by solving the subspace-averaging problem:

$$\mathbb{P} \mapsto \begin{cases} \max\limits_{\mathbf{x}} \ \mathbf{x}^{\dagger}\Xi\mathbf{x} \\ \text{s.t. } \|\mathbf{U}_2\mathbf{x}\|^2 \leq P_R. \end{cases} \tag{A-64}$$

Appendix B

Appendix Math

In this part, we introduce the mathematical tools which are much used in this thesis, and they are matrix manipulation, convex optimization and robust optimization. For convex optimization, we review the basic concepts such as the property of convex set, Euclidean balls and ellipsoids. Finally, from various types of convex optimization problems, we focus on SDP, SOCP problems for further reference of later chapters. General conic data uncertainty type of optimization is then discussed in the context of robust optimization, together with basic stochastically robust solutions for data uncertainty problems. The materials regarding matrix analysis in this chapter are largely based on [39].

B.1 Matrix Manipulation

First, we introduce some frequently used matrix definitions in our thesis.

Definition 1. *Given a matrix* $\mathbf{A} \in \mathcal{C}^{n\times n}$*, it is said to be a non-singular matrix if*

$$\mathbf{A}\mathbf{v} = 0, \text{ only when } \mathbf{v} = 0. \tag{A-1}$$

Then $\mathbf{A}$ *is a singular matrix, if*

$$\exists \mathbf{v} \neq 0, \text{ s.t. } \mathbf{A}\mathbf{v} = 0. \tag{A-2}$$

Definition 2. *A matrix* $\mathbf{A} \in \mathcal{C}^{n\times n}$ *is said to be a Hermitian matrix if* $\mathbf{A} = \mathbf{A}^*$.

Definition 3. *A matrix* $\mathbf{A} \in \mathcal{C}^{n\times n}$ *is a positive definite matrix, if*

$$\mathbf{v}^*\mathbf{A}\mathbf{v} \geq 0 \tag{A-3}$$

for any non-zero complex vector $\mathbf{v}$*, where* $\mathbf{v}^*$ *denotes the conjugate transpose of* $\mathbf{v}$*. Note that the quantity* $\mathbf{v}^*\mathbf{A}\mathbf{v}$ *is always real because* $\mathbf{A}$ *is a Hermitian matrix.*

B.2 Eigenvalue and Eigenvector

For almost all vectors, their directions are changed once multiplied by a matrix $\mathbf{A}$, but there are some special vectors that stay their original directions. We define these vectors as follows.

Given an $n \times n$ matrix $\mathbf{A}$ with entries in field $\mathcal{F}$ (real or complex), any vector $\mathbf{v}$ that satisfies

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}, \text{ for some } \lambda, \tag{A-4}$$

is called the eigenvector of $\mathbf{A}$, and the corresponding λ is the eigenvalue. The eigenvalue set of $\mathbf{A}$ is called the spectrum of $\mathbf{A}$. When $\mathbf{A}$ is square, the eigenvalue and eigenvector of $\mathbf{A}^2$ can be found as

$$\mathbf{A}^2\mathbf{v} = \mathbf{A}\lambda\mathbf{v} = \lambda\mathbf{A}\mathbf{v} = \lambda^2\mathbf{v}. \tag{A-5}$$

That is, for the matrix $\mathbf{A}^k$, the eigenvector stays the same but the eigenvalue becomes λ^k.

Of field $\mathcal{F}^n$, a subspace $\mathcal{S}$ is a closed subset which satisfies:

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{S}, \forall a, b \in \mathbf{F} : (a\mathbf{x} + b\mathbf{y}) \in \mathcal{S}. \tag{A-6}$$

A vector set $\{\mathbf{x}_1, \ldots, \mathbf{x}_l\}$ with elements of v, and that every element of $\mathcal{S}$ can be written as a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_l$ is a spanning set of $\mathcal{S}$. In other words, $\mathcal{S}$ is the column space of matrix $\{\mathbf{x}_1, \ldots, \mathbf{x}_l\}$. The kernel of a matrix $\mathbf{B}$ represented as $\mathsf{ker}(\mathbf{B})$ is the subspace spanned by vectors $\mathbf{x}$ that satisfies $\mathbf{B}\mathbf{x} = 0$, which is the formed by the eigenvectors of $\mathbf{B}$.

Given two matrices of $\mathcal{C}^{n\times n}$, $\mathbf{A}$ and $\mathbf{B}$, if there exists $(\lambda, \mathbf{v})$ that

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{B}\mathbf{v}, \tag{A-7}$$

then we say that $(\lambda, \mathbf{v})$ is an eigen-pair of the pencil $(\mathbf{A}, \mathbf{B})$. The generalised eigenvalue problem is exactly finding the eigen-pair of a matrix pencil.

B.3 Matrix Diagonalisation and Similar Matrices

For matrix $\mathbf{A}$ with rank r, and the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$, the eigenvector matrix of $\mathbf{A}$ is defined as $\mathbf{V} = (\mathbf{v}_1 \ldots \mathbf{v}_r)$. We can get a diagonal eigenvalue matrix $\mathbf{\Sigma} = \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$.

Definition 4. *Let* $\mathbf{S}$ *be any invertible matrix. Then matrix* $\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}$ *is similar to matrix* $\mathbf{A}$.

Also, we can use eigenvalue to detect the characteristic of a matrix. A symmetric matrix is positive semidefinite if all its eigenvalues are nonnegative: $\mathbf{A} \succeq 0$. Similarly, $\mathbf{A} \succ 0$, $\mathbf{A} \prec 0$ and $\mathbf{A} \preceq 0$ means that $\mathbf{A}$ is positive definite, negative semidefinite, and negative definite, respectively.

B.4 Hadamard Product

Given two matrices $\mathbf{A} \in \mathcal{R}^{n\times n}$ and $\mathbf{B} \in \mathcal{R}^{n\times n}$, the Hadamard product is defined as

$$(\mathbf{A} \odot \mathbf{B})_{ij} = (\mathbf{A}_{ij}) \times (\mathbf{B})_{ij}, \text{ for } i, j = 1, \ldots, n. \tag{A-8}$$

B.5 Kronecker Product

Given two matrices $\mathbf{A} \in \mathcal{R}^{m\times n}$ and $\mathbf{B} \in \mathcal{R}^{p\times q}$, their Kronecker product, denoted as $\mathbf{A} \otimes \mathbf{B}$, is

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \cdots & a_{1n}\mathbf{Q} \\ \vdots & \ldots & \vdots \\ a_{n1}\mathbf{B} & \cdots & a_{nn}\mathbf{Q} \end{bmatrix}. \tag{A-9}$$

B.6 Singular Value Decomposition (SVD)

Given a matrix $\mathbf{A} \in \mathcal{C}^{M\times N}$, $M \geq N$ and $r \triangleq \mathsf{rank}(\mathbf{A}) \leq \min(M, N)$, for two sets of vectors $\mathbf{u}$'s which are the eigenvectors of $\mathbf{A}\mathbf{A}^*$ and $\mathbf{v}$'s the eigenvectors of $\mathbf{A}^*\mathbf{A}$, as $\mathbf{A}\mathbf{A}^*$ and $\mathbf{A}^*\mathbf{A}$ are symmetric matrices, we can choose $\mathbf{u}$'s and $\mathbf{v}$'s to be orthogonal vectors. As $\mathbf{A}(\mathbf{A}\mathbf{A}^*) = (\mathbf{A}^*\mathbf{A})\mathbf{A}$, we have

$$\mathbf{A}\mathbf{v}_1 = \sigma_1\mathbf{u}_1, \ldots, \mathbf{A}\mathbf{v}_r = \sigma_r\mathbf{u}_r. \tag{A-10}$$

Accordingly, we have

$$\mathbf{A}\left(\mathbf{v}_1 \cdots \mathbf{v}_r\right) = \left(\mathbf{u}_1 \cdots \mathbf{u}_r\right) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}. \tag{A-11}$$

Besides, for the null-space of $\mathbf{A}$, denoted as $\mathbb{N}(\mathbf{A})$, which has the $n - r$ eigenvectors $\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n$, and for the null-space of $\mathbf{A}^T$ $\mathbb{N}(\mathbf{A}^T)$, which has the $m - r$ eigenvectors $\mathbf{u}_{r+1}, \ldots, \mathbf{u}_{m-r}$, we define

$$\mathbf{V} = \left(\mathbf{v}_1 \cdots \mathbf{v}_n\right), \text{ and } \mathbf{U} = \left(\mathbf{v}_1 \cdots \mathbf{v}_m\right), \tag{A-12}$$

and $\mathbf{V}$ and $\mathbf{U}$ are unitary matrices. We have the SVD of $\mathbf{A}$ as

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger, \tag{A-13}$$

where $\mathbf{\Sigma}$ is a diagonal matrix with diagonal entries $\sigma_n \geq \cdots \geq \sigma_1 \geq 0$ which are called the singular values of $\mathbf{A}$. Also, $\mathbf{U}$,$\mathbf{V}$ are referred to as the left and right singular vectors of $\mathbf{A}$.

B.7 Commutation Matrix

Given a matrix $\mathbf{A} \in \mathcal{R}^{m\times n}$, define the commutation matrix of $\mathbf{A}$ as $\mathbf{K} \in \mathcal{R}^{mn\times mn}$

$$\mathbf{K}\mathsf{vec}(\mathbf{A}) = \mathsf{vec}(\mathbf{A}^T), \tag{A-14}$$

where

$$\mathsf{vec}(\mathbf{A}) = (\mathbf{A}_{11}, \ldots, \mathbf{A}_{m1}, \ldots, \mathbf{A}_{1n}, \ldots, \mathbf{A}_{mn})^T \in \mathcal{R}^{mn\times 1}, \tag{A-15}$$

with $\mathbf{A}_{i,j}$, for $i \in (1,m), j \in (1,n)$, gives the (i,j)th entry of the matrix $\mathbf{A}$. Moreover,

$$\mathbf{K} = \sum_{i=1}^{m}\sum_{j=1}^{n}(\mathbf{H}_{ij} \otimes \mathbf{H}_{ij}^{'}), \tag{A-16}$$

where $\mathbf{H}_{ij} \in \mathcal{R}^{m\times n}$ has the (i,j)th entry being 1 with other entries being 0's.

B.8 Rayleigh Quotient

Given a symmetric matrix $\mathbf{A}$ and a positive definite matrix $\mathbf{B}$, and that $\lambda_1 \leq \cdots \leq \lambda_n$ being the eigenvalues of pencil $(\mathbf{A},\mathbf{B})$, the generalised Rayleigh quotient problem is defined on the set of all $n \times p$ full-rank matrices as:

$$f(\mathbf{x}) = \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{B}\mathbf{x}}. \tag{A-17}$$

Defining $\mathbf{x}^*$ is a global minimiser of (A-17), then we have:

(i) $\mathsf{span}(\mathbf{x}^*)$ is the leftmost invariant subspace of $(\mathbf{A},\mathbf{B})$;

(ii) $f(\mathbf{x}) = \sum_{i=1}^{p}\lambda_i$.

B.9 Mathematical Optimization

A mathematical optimization problem is a problem with a well-defined objective and a list of well-defined constraints in mathematical form. We use an example below to introduce it in details. Consider a mathematical optimization problem, which has the form:

$$\begin{aligned} &\min_{\mathbf{x}} \ f_0(\mathbf{x}) \\ &\text{s.t.} \ \begin{cases} f_i(\mathbf{x}) \leq b_i, \ \text{for } i = 1,\ldots,m \\ h_j(\mathbf{x}) = 0, \ \text{for } j = 1,\ldots,n, \end{cases} \end{aligned} \tag{A-18}$$

where the vector $\mathbf{x}$ contains the unknown decision variables of the optimization problem, $f_0(\mathbf{x})$ is the objective function which maps the variable $\mathbf{x} \in \mathcal{R}^n$ to $f_0(\mathbf{x}) \in \mathcal{R}$, the functions $f_i(\mathbf{x}) : \mathcal{R}^n$ are the inequality constraint functions, and $b_1,\ldots,b_m$ are constant values which are the limits, or bounds, for the constraints. Likewise, the functions $h_j(\mathbf{x}) : \mathcal{R}^n \mapsto \mathcal{R}$, for $j = 1,\ldots,n$, are the equality constraint functions. If there exist some $\mathbf{x}$ that satisfies all the constraints, then we say that the problem is feasible; otherwise, it is infeasible. A vector $\mathbf{x}^*$ is called the optimal solution of (A-18), if it has the smallest objective value among all those vectors that satisfy the constraints. Mathematically, that is, for any $\mathbf{z}$ with $f_1(\mathbf{z}) \leq b_z,\ldots,f_m(\mathbf{z}) \leq b_m, h_1(\mathbf{z}) = 0,\ldots,h_n(\mathbf{z}) = 0$, we have $f_0(\mathbf{z}) \geq f_0(\mathbf{x}^*)$.

Generally, the optimization problems can be very difficult to solve or compute, even if solvable. However, a class of optimization called convex optimization is proved to be solvable and easy to compute the optimal solution. In this thesis, motivated by the power of convex optimization, we always seek to convert our optimization problems in the TWRC into some forms of convex optimization problems, which can lead to efficient computation of the global optimal solutions.

A convex optimization function as shown in Fig. B.1 is the one in which the objective and constraint functions are all convex, which means that they satisfy the inequality:

$$f_i(\alpha\mathbf{x} + \beta\mathbf{y}) \geq \alpha f_i(\mathbf{x}) + \beta f_i(\mathbf{y}) \tag{A-19}$$

for all $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$ and all $\alpha, \beta \in \mathcal{R}$ with $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta \geq 0$.

Figure B.1: An illustration of a convex function.

Besides the methods described in (A-19), the convexity of a function $f(x)$ can also be checked by the first order and second order conditions. This is described in the following.

- **First order condition**—For a differentiable function $f(\mathbf{x})$, whose gradient $\nabla f(\mathbf{x})$ exists for all $\mathbf{x} \in \mathsf{dom}(f)$, then $f(x)$ is convex if and only if $\mathsf{don}(f)$ is a convex set, and for all $\mathbf{x}_1$ and $\mathbf{x}_2 \in \mathsf{dom}(f)$, the following inequality holds true:

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + \nabla f(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1). \tag{A-20}$$

- **Second order condition**—For a twice differentiable function f, if its second derivative $\nabla^2 f(\mathbf{x})$ exists and satisfies that:

$$\nabla^2 f(\mathbf{x}) \geq 0, \tag{A-21}$$

  then $f(\mathbf{x})$ is a convex function. For a function $f(\mathbf{x})$ with $\mathbf{x} \in \mathcal{R}^n$, it is equivalent to $f^{''}(x) \geq 0$, which means that the derivative of $f(\mathbf{x})$ is either constant or increasing.

The convexity property can make optimization in some sense "easier" than the general case. According to its convexity, any local minimum must be a global minimum. Once a problem is proved to be convex or can be recast into a convex problem, it can be directly solved by existing convex optimization tools, such as SEDUMI or CVX. For the rest of this section, we describe some important examples of convex sets which we will encounter in this thesis.

B.10 Lagrange Duality Problem

Regarding the original optimization problem as the primary problem, we can formulate a dual problem of the primary one. Generally, the solution of the dual problem and that of the primary problem would not be the same. To find the dual of an optimization problem, we look at the optimization problem from

a different perspective. Particularly, the aim is to seek a lower bound of the original problem. However, for convex optimization problems, it is proved that the solution of the dual problem is also the solution of the primary problem. For the optimization problem (A-18), the Lagrangian function is defined as

$$\Lambda(\mathbf{x}, \lambda, \nu) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{j=1}^{n} \nu_j h_j(\mathbf{x}), \tag{A-22}$$

where $\mathbf{\Lambda} = (\lambda_1, \dots)$ and $\boldsymbol{\nu} = (\nu_1, \dots)$ are called the dual variables or Lagrange multiplier vectors of the Lagrange function. We define the Lagrange dual function g as

$$g(\lambda, \nu) = \inf_{\mathbf{x} \in \mathbb{D}} \left\{ f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{j=1}^{n} \nu_j h_j(\mathbf{x}) \right\}. \tag{A-23}$$

From (A-23), we can see that the Lagrange dual function is a concave function, and assuming that p^{opt} is the optimal solution of the primary function (A-18), and p^* is the solution of (A-23), and with $\mathbf{\Lambda} \geq \mathbf{0}$ and $\boldsymbol{\nu} \geq \mathbf{0}$, we always have

$$p^* \leq p^{\mathsf{opt}}.$$

Now, we introduce the *Slater's condition*, which says that: The difference between the dual optimum and the primary optimum is called the duality gap. If there exists $\mathbf{x} \in \mathsf{dom} = \bigcap_{i=0}^{m} f(i)$ that satisfies the constraints of (A-18), then there's a 0 duality dap between the primary and dual problems, which means that the value of the primal and dual problems are equal. This is called the strong duality.

B.11 Euclidean Balls and Ellipsoids

A (Euclidean) ball (or just ball) in $\mathcal{R}^n$ is of the form

$$B(\mathbf{x}_c, r) = \{\mathbf{x} \mid \|\mathbf{x} - \mathbf{x}_c\| \leq r\} = \{\mathbf{x} \,|(\mathbf{x} - \mathbf{x}_c)^T(\mathbf{x} - \mathbf{x}_c) \leq r^2\}, \tag{A-24}$$

where $r \geq 0$, and $\|\cdot\|_2$ denotes the Euclidean norm, i.e., $\|\mathbf{u}\|_2 = (\mathbf{u}^T\mathbf{u})^{1/2}$.

The vector $\mathbf{x}_c$ is the centre of the ball and the scalar r is its radius; $B(\mathbf{x}_c, r)$ consists of all points within a distance r from the center $\mathbf{x}_c$. Another common representation for the Euclidean ball is:

$$B(\mathbf{x}_c, r) = \{\mathbf{x}_c + r\mathbf{u} \mid \|\mathbf{u}\| \leq 1\}, \tag{A-25}$$

and a Euclidean ball is a convex set.

A related family of convex sets is the ellipsoids which has the form

$$\varepsilon = \{\mathbf{x} | (\mathbf{x} - \mathbf{x}_c)^T \mathbf{P}^{-1} (\mathbf{x} - \mathbf{x}_c) \leq 1\}, \tag{A-26}$$

where $\mathbf{P} = \mathbf{P}^T \succeq \mathbf{0}$ is symmetric and positive definite. Again, $\mathbf{x}_c \in \mathcal{R}^n$ is the center of the ellipsoid. The matrix $\mathbf{P}$ determines how far the ellipsoid extends in every direction from $\mathbf{x}_c$; the lengths of the semi-axes of ε are given by $\sqrt{\lambda_i}$, where λ_i are the eigenvalues of $\mathbf{P}$. A ball is an ellipsoid with $\mathbf{P} = r^2\mathbf{I}$.

Another common representation of an ellipsoid is

$$\varepsilon = \{\mathbf{x}_c + \mathbf{A}\mathbf{u},\ |\ \forall \|\mathbf{u}\| \leq 1\}. \tag{A-27}$$

where $\mathbf{A}$ is square and nonsingular. In this representation, we can assume, without loss of generality, that $\mathbf{A}$ is symmetric and positive definite. By taking $\mathbf{A} = \mathbf{P}^{1/2}$, this representation gives (A-27).

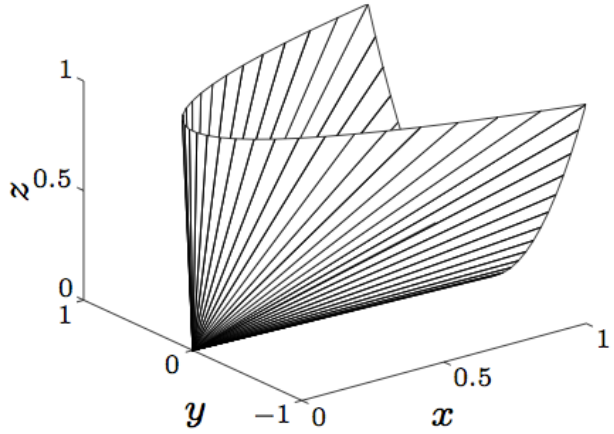

Figure B.2: The positive semidefinite cone.

B.12 The Positive Semidefinite Cone

The set of all symmetric positive semidefinite matrices of particular dimension is called the positive semidefinite cone. It can be formed by intersection of an infinite number of half-spaces in the vectorised variable matrix (as shown in the figure), each half-space having partial boundary containing the origin in an isomorphic subspace. Hence the positive semidefinite cone is convex. It is a unique immutable proper cone in the ambient space of symmetric matrices. The positive definite (full-rank) matrices comprise the cone interior, while all singular positive semidefinite matrices (having at least one 0 eigenvalue) reside on the cone boundary. The only symmetric positive semidefinite matrix having all eigenvalues resides at the origin. In low dimension, the positive semidefinite cone is shown to be a circular cone by way of an isometric isomorphism relating matrix space to vector space.

For a 2×2 symmetric matrix such as

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathcal{S}_+^2, \tag{A-28}$$

the positive semidefinite cone is as shown in Fig. B.2. In one dimension with 1×1 symmetric matrices, the nonnegative ray is a circular cone.

B.13 Examples for Quadratic Convex optimization Problem

We use the definition of quadratic program (QP) from [14]. A quadratic program is a convex optimization problem with a quadratic convex objective function and linear constrained functions. The standard format of a QP is expressed as

$$\begin{aligned} \min_{\mathbf{x}} \ & \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r \\ \text{s.t.} \ & \begin{cases} \mathbf{G}\mathbf{x} \leq \mathbf{h}, \\ \mathbf{A}\mathbf{x} = \mathbf{b}, \end{cases} \end{aligned} \tag{A-29}$$

where $\mathbf{P} \in \mathcal{S}_+^n$, $\mathbf{G} \in \mathcal{R}^{m\times n}$, and $\mathbf{A} \in \mathcal{R}^{p\times n}$. If the objective function, the inequality constrained functions are all convex quadratic functions, then we call this type of problem the quadratically constrained quadratic program (QCQP). The general QCQP form is

$$\begin{aligned} \min_{\mathbf{x}} \ & \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T x + r \\ \text{s.t.} \ & \begin{cases} \frac{1}{2}\mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i^T x + r_i \leq 0, \text{ for } i = 1, \ldots, m, \\ \mathbf{A}\mathbf{x} = \mathbf{b}, \end{cases} \end{aligned} \tag{A-30}$$

where $\mathbf{P}_i \in \mathcal{S}_+^n$, for $i = 0, 1, \ldots, m$.

Now, we introduce SOCP and SDP, the two types of problems that are closely related to QCQP problems and that our optimization problems in TWRC take such forms (see later chapters).

B.14 SOCP

Consider a typical form of an SOCP problem as :

$$\begin{aligned} \min_{\mathbf{x}} \ & \mathbf{f}^T(\mathbf{x}) \\ \text{s.t.} \ & \begin{cases} \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\| \leq \mathbf{c}_i^T + d_i, \text{ for } i = 1, \ldots, m, \\ F(\mathbf{x}) = g, \end{cases} \end{aligned} \tag{A-31}$$

where $\mathbf{f} \in \mathcal{R}^n$, $\mathbf{A}_i \in \mathcal{R}^{n_i\times n}$, $\mathbf{b}_i \in \mathcal{R}^{n_i}$, $\mathbf{c}_i \in \mathcal{R}^n$, and $d_i \in \mathcal{R}$, with $\mathbf{x}$ being the optimising variable. Also, $\mathbf{f}^T(x)$ is a convex function and $F(\mathbf{x})$ is a linear function of $\mathbf{x}$. Such problem is considered as SOCP since it has second-order cone constraints. When $\mathbf{A}_i = 0$ for $i = 1, \ldots, m$, SOCP is reduced to a general linear program. When $\mathbf{c}_i = 0$ for $i = 1, \ldots, m$, SOCP is equivalent to a convex QCQP. As SOCP constraints can be written as LMIs, the problem can be reformulated as an instance of SDP. SOCPs can be solved with great efficiency by interior point methods.

B.15 SDP

SDP is a subfield of convex optimization concerned with the optimization of a linear objective function with linear inequality constraints, and a non-negative matrix constraint. The constraints of a standard SDP are the intersection of the cone of positive semidefinite matrices with an affine space, i.e., a spectrahedron. SDP is a relatively new field of optimization. Besides its convexity property, it is of growing

interest for several reasons. In reality, many practical problems in field such as engineering, control, and combinatorial optimization, can be formulated or approximated as SDP problems. In automatic control theory, SDPs are used in the context of LMIs. SDPs are in fact a special case of cone programming and can be efficiently solved by interior point methods. All linear programs can be reformulated as SDPs. Also, using the hierarchies of SDPs, the solutions of polynomial optimization problems can be approximated. Also, SDP has been used in the optimization for numerous complex systems.

Denote $\mathcal{S}^N$ as the space of all $N \times N$ real symmetric matrices. The space is equipped with the inner product (where trace denotes the trace function):

$$\langle \mathbf{A}, \mathbf{B} \rangle = \mathsf{trace}(\mathbf{A}^T \mathbf{B}) = \sum_{i,j=1}^{n} \mathbf{A}_{ij} \mathbf{B}_{ij}. \tag{A-32}$$

Denote by $\mathcal{S}_+^N$ the convex cone of positive semidefinite $N \times N$ matrices. This cone defines a partial order for $\mathbf{A}, \mathbf{B} \in \mathcal{S}^N$ by $\mathbf{A} \succeq \mathbf{B}$ whenever $\mathbf{A} - \mathbf{B}$ is positive semidefinite.

Linear SDP deals with optimization problems of the type:

$$\begin{aligned} \min_{\mathbf{X} \in \mathcal{S}^N} \ & \mathsf{trace}(\mathbf{C}^T \mathbf{X}) \\ \text{s.t.} \ & \begin{cases} \mathsf{trace}(\mathbf{A}_i^T \mathbf{X}) = b_i, \ \text{for } i = 1, \ldots, m, \\ \mathbf{X} \succeq \mathbf{0}. \end{cases} \end{aligned} \tag{A-33}$$

We refer to this as the primal SDP. Similar to linear programming, we can have a dual SDP:

$$\begin{aligned} \max_{\mathbf{y} \in \mathcal{R}^m} \ & \mathbf{b}^T \mathbf{y} \\ \text{s.t.} \ & \sum_{i=1}^{m} \mathbf{y}_i \mathbf{A}_i \preceq \mathbf{C}. \end{aligned} \tag{A-34}$$

For convenience, an SDP will often be presented in a slightly different but equivalent form. For example, linear expressions involving nonnegative scalar variables can be added to the SDP specification and it remains as an SDP because each variable can be incorporated into the matrix $\mathbf{X}$ as a diagonal entry ($\mathbf{X}_{ii}$ for some i). To ensure that, constraints $\mathbf{X}_{ij} = 0$ can be added for all. As another example, note that for any positive semidefinite matrix $\mathbf{X}$, there exists a set of vectors $\mathbf{v}_i$ such that the (i, j)th entry of $\mathbf{X}$ is $\mathbf{X}_{ij} = (\mathbf{v}_i, \mathbf{v}_j)$ the scalar product of $\mathbf{v}_i$ and $\mathbf{v}_j$. Therefore, SDPs are often formulated in terms of linear expressions on scalar products of vectors. Given the solution to the SDP in the standard form, the vectors $\mathbf{v}_i$ can be recovered in $O(n^3)$ time (e.g., by using a Cholesky decomposition of $\mathbf{X}$).

The strong duality mentioned above holds true for all linear programs; however, not every SDP satisfies the strong duality. In general, the value of the dual may lie strictly below that of the primal.

(i) Suppose the primal SDP problem (A-33) is lower bounded and strictly feasible. Then there is an optimal solution $\mathbf{y}^*$ to the dual and

$$\mathsf{trace}(\mathbf{C}^T \mathbf{X}) = \mathbf{b}^T \mathbf{y}. \tag{A-35}$$

(ii) Suppose the dual problem is upper bounded and strictly feasible. Then there is an optimal solution $\mathbf{X}^*$ to the primal SDP problem (A-33) and (i) holds the equality.

The weak duality theorem of SDP is that the value of the primal SDP is at least the value of the dual SDP. Therefore, any feasible solution to the dual SDP is a lower bound of the primal SDP value, and equivalently, the feasible solution to the primal SDP is an upper-bound of the dual SDP value. As such,

$$\mathsf{trace}(\mathbf{C}^T\mathbf{X}) - \mathbf{b}^T\mathbf{y} \geq 0, \tag{A-36}$$

where the inequality holds because both of the matrices are positive semidefinite.

B.16 Optimization with Uncertainty

When we use mathematic methods to model a system, there would be uncertainty of certain parameters. For example, in wireless communication system, there is fluctuation of channel coefficients during transmission, so the channel factor may not be known exactly for network optimization. The uncertainty in the parameters will dramatically change the final optimal solution.

There are two typical methods for dealing with uncertainty. One tackles the case when the uncertainty factors have bounded perturbation. Robust optimization provides robustness in the optimization in these situations. Another one is stochastic optimization which is based on probability distribution of the uncertainty. Linear and conic data uncertainty in optimization has been studied in [9, 11].

B.17 Robust Optimization

The uncertainty of data can be characterised by a bounded uncertainty data set. Optimization of these problems is to find the optimal results for the worst case. Suppose that we have a linear programming optimization problem which contains the uncertain linear inequality:

$$\min_{\mathbf{x}} \quad \mathbf{c}^T\mathbf{x} + d \tag{A-37}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{b}, \tag{A-38}$$

where we have $\mathbf{c} \in \mathcal{R}^n$, $d \in \mathcal{R}$, $\mathbf{b} \in \mathcal{R}^m$ and $\mathbf{A} \in \mathcal{R}^{m\times n}$ belong to the uncertainty domain, which is defined by a set $\mathcal{U}$ through the perturbation set $\mathcal{Z}$:

$$\mathcal{U} = \left\{ \left[\begin{array}{c|c} \mathbf{c}^T & d \\ \hline \mathbf{A} & \mathbf{b} \end{array}\right] = \underbrace{\left[\begin{array}{c|c} \mathbf{c}_0^T & d_0 \\ \hline \mathbf{A}_0 & \mathbf{b}_0 \end{array}\right]}_{\text{nominal data } \mathbf{D}_0} + \sum_{l=1}^{L} \zeta_l \underbrace{\left[\begin{array}{c|c} \mathbf{c}_l^T & d_l \\ \hline \mathbf{A}_l & \mathbf{b}_l \end{array}\right]}_{\text{basic shift } \mathbf{D}_l} : \zeta \in \mathcal{Z} \subset \mathcal{R}^L \right\}. \tag{A-39}$$

Within the uncertainty domain, we can see that although the uncertainty of d affects the optimal results, but has no influence on the optimal solutions, so we can ignore d in the remaining discussion below. Considering these uncertainty, we can write a robust counterpart of the optimization problem

(A-37) as

$$\begin{aligned} &\min_{\mathbf{x}} \ \mathbf{c}^T\mathbf{x} \\ &\text{s.t.} \ \begin{cases} \mathbf{A}\mathbf{x} \leq \mathbf{b}, \\ \forall(\mathbf{c},\mathbf{A},\mathbf{b}) \in \mathcal{U}. \end{cases} \end{aligned} \tag{A-40}$$

B.18 Conic Uncertainty Set

In this thesis, we focus on uncertain conic problems. Consider a conic optimization problem as:

$$\min_{\mathbf{x}} \quad \mathbf{c}^T\mathbf{x} + d \tag{A-41}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{K} \tag{A-42}$$

where $\mathbf{x} \in \mathcal{R}^n$ is the optimising variable and $\mathbf{K}$ is a closed convex cone, which could be:

- A positive orphan $\mathcal{R}^m_+$, which is of the form $\{\mathbf{a}_i^T\mathbf{x} - \mathbf{b}_i \geq 0, 1 \leq i \leq m\}$;
- A Lorentz (second-order cone), which is of the form $\{\|\mathbf{A}_i\mathbf{x} - \mathbf{b}_i\| \leq \mathbf{c}_i^T\mathbf{x} - d_i, 1 \leq i \leq m\}$;
- A semidefinitite cone $\mathcal{S}^k_+$, in the form of $\{\mathbf{A}_i\mathbf{x} - \mathbf{B}_i \geq 0, 1 \leq i \leq m\}$.

The uncertain conic problem of (A-41) is the problem with fixed structure and uncertain data parameterised by a perturbation vector $\zeta \in \mathcal{R}^L$ through a known perturbation set $\mathcal{Z} \in \mathcal{R}^L$

$$(\mathbf{c}, d, \{\mathbf{A}_i, \mathbf{b}_i\}_{i=1}^m) = \left(\mathbf{c}^0, d^0, \{\mathbf{A}_i^0, \mathbf{b}_i^0\}_{i=1}^m\right) + \sum_{l=1}^{L} \zeta_l \left(\mathbf{c}^l, d^l, \{\mathbf{A}_i^l, \mathbf{b}_i^l\}_{i=1}^m\right). \tag{A-43}$$

B.19 The Worst-Case Model

Here we present a few important lemmas which facilitate robust optimization.

Lemma 1. S-Lemma 1: *Letting Hermitian matrices $\mathbf{A}_j \in \mathcal{C}^{n\times n}$, vectors $\mathbf{b}_j \in \mathcal{C}^n$, and scalars $c_j \in \mathcal{C}$, we define the following functions: $f_j(\mathbf{x}) = \mathbf{x}^H\mathbf{A}_j\mathbf{x} + 2\mathrm{R}(\mathbf{b}^H\mathbf{x}) + c_j$, for $j = 0, 1, 2, \ldots, n$. Then the following 2 conditions are the same:*

1. *$f_0(\mathbf{x}) \geq 0$ for every $\mathbf{x} \in \mathcal{C}^n$ such that $f_i(\mathbf{x}) \succeq 0$, for $i = 1, \ldots, n$;*
2. *There exists $\lambda_i \geq 0$, for $i = 1, \ldots, n$, and $\begin{pmatrix} \mathbf{A}_0 & \mathbf{b}_0 \\ \mathbf{b}_0^H & c_0 \end{pmatrix} \geq \sum_{i=1}^{i=n} \lambda_i \begin{pmatrix} \mathbf{A}_i & \mathbf{b}_i \\ \mathbf{b}_i^H & c_i \end{pmatrix} \succeq 0$.*

Lemma 2. S Lemma 2: *Let $\mathbf{P}$ be a symmetric matrix and $\mathbf{A}$ be a rectangle matrix. Then*

$$\mathbf{P} - \mathbf{A}^T\mathbf{A} \succeq 0 \tag{A-44}$$

if and only if

$$\begin{bmatrix} \mathbf{P} & \mathbf{A}^T \\ \mathbf{A} & \mathbf{I} \end{bmatrix} \succeq 0. \tag{A-45}$$

In matrix theory, the definition of Schur component of a block matrix

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{P}^T & \mathbf{Q}^T \\ \hline \mathbf{Q} & \mathbf{R} \end{array}\right] \tag{A-46}$$

is $\mathbf{P} - \mathbf{Q}^T\mathbf{R}^{-1}\mathbf{A}$, and we have Lemma 3 as follows.

Lemma 3. Schur Component Lemma: *Let* $\mathbf{A}$ *be a symmetric block matrix*

$$\mathbf{A} = \left[\begin{array}{c|c} \mathbf{P} & \mathbf{Q}^T \\ \hline \mathbf{Q} & \mathbf{R} \end{array}\right]. \tag{A-47}$$

Then $\mathbf{R} \succeq 0$ *if and only if*

$$\mathbf{P} - \mathbf{Q}^T\mathbf{R}^{-1}\mathbf{Q} \succeq 0. \tag{A-48}$$

Proof. This result can be proved by performing Gaussian elimination on matrix $\mathbf{A}$ by right multiplexing a lower triangle matrix, i.e.,

$$\begin{aligned} \mathbf{AL} &= \left[\begin{array}{c|c} \mathbf{P}^T & \mathbf{Q}^T \\ \hline \mathbf{Q} & \mathbf{R} \end{array}\right] \left[\begin{array}{c|c} \mathbf{I}_p & 0 \\ \hline -\mathbf{R}^{-1}\mathbf{C} & \mathbf{I}_q \end{array}\right] \\ &= \left[\begin{array}{c|c} \mathbf{P} - \mathbf{Q}^T\mathbf{R}^{-1}\mathbf{Q} & \mathbf{Q}^T \\ \hline 0 & \mathbf{R} \end{array}\right] \\ &= \left[\begin{array}{c|c} \mathbf{I}_p & \mathbf{Q}^T\mathbf{R}^{-1} \\ \hline 0 & \mathbf{I}_q \end{array}\right] \left[\begin{array}{c|c} \mathbf{P} - \mathbf{Q}^T\mathbf{R}^{-1}\mathbf{Q} & 0 \\ \hline 0 & \mathbf{R} \end{array}\right]. \end{aligned} \tag{A-49}$$

Also, we have $\mathbf{A} = \mathbf{A}^T \succeq 0$. Therefore, we also get

$$\begin{aligned} &\Longleftrightarrow \mathbf{u}^T\mathbf{P}\mathbf{u} + 2\mathbf{u}^T\mathbf{Q}^T\mathbf{v} + \mathbf{v}^T\mathbf{R}\mathbf{v}, \forall \mathbf{u}, \mathbf{v} \\ &\Longleftrightarrow \forall \mathbf{u} : 0 \leq \min_{\mathbf{v}}\{\mathbf{u}^T\mathbf{P}\mathbf{u} + 2\mathbf{u}^T\mathbf{Q}^T\mathbf{v} + \mathbf{v}^T\mathbf{R}\mathbf{v}\} = \mathbf{u}^T\mathbf{P}\mathbf{u} + 2\mathbf{u}^T\mathbf{Q}^T\mathbf{R}^{-1}\mathbf{Q}\mathbf{u}. \end{aligned} \tag{A-50}$$

□

Now, consider the uncertainty QCQP:

$$\mathbf{P} \mapsto \begin{cases} \min\limits_{\mathbf{x}} \ \mathbf{C}^T\mathbf{x} \\ \quad \text{s.t.} \ \begin{cases} \mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} - 2\mathbf{b}^T\mathbf{x} \leq \Gamma \\ \forall (\mathbf{A}, \mathbf{b}, \Gamma) \in \mathbf{U}, \end{cases} \end{cases} \tag{A-51}$$

where $\mathbf{U} = \{(\mathbf{A}, \mathbf{b}, \Gamma) = (\mathbf{A}_0, \mathbf{b}_0, \Gamma_0) + \sum_{k=1}^{N} \mathbf{u}_k(\mathbf{A}_k, \mathbf{b}_k, \Gamma_k) | \|\mathbf{u}\| \leq 1\}$. Let:

$$\mathbf{F}(\mathbf{x}) = (\mathbf{A}_0\mathbf{x}, \mathbf{A}_1\mathbf{x}, \ldots, \mathbf{A}_k\mathbf{x}). \tag{A-52}$$

The constraint can be rewritten as (A-53):

$$
\begin{aligned}
&-\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}+2\mathbf{b}^T\mathbf{x}+\Gamma= \\
&\begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix}^T \left(\begin{pmatrix} \Gamma_0+2\mathbf{x}^T\mathbf{b}_0 & \frac{\Gamma_1}{2}+\mathbf{x}^T\mathbf{b}_1 & \dots & \frac{\Gamma_k}{2}+\mathbf{x}^T\mathbf{b}_k \\ \frac{\Gamma_1}{2}+\mathbf{x}^T\mathbf{b}_1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \frac{\Gamma_k}{2}+\mathbf{x}^T\mathbf{b}_k & 0 & \dots & 0 \end{pmatrix} - \mathbf{F}(\mathbf{x})^T\mathbf{F}(\mathbf{x}) \right) \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix} \geq 0, \\
&\qquad\qquad \forall \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & -\mathbf{I} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{u} \end{pmatrix} \geq 0. \qquad \text{(A-53)}
\end{aligned}
$$

B.20 Statistical Model

For stochastic uncertainty, the uncertainty is random, but follows a certain probability distribution. In the simplest case, the pdf is considered known. If it is only partially known, this will be much more difficult. The stochastic optimization model for (A-37) can be constructed in the form as:

$$
\begin{aligned}
&\min_{\mathbf{x}} \ \mathbf{c}^T\mathbf{x} \\
&\text{s.t. } \mathsf{Prob}\{\mathbf{A}\mathbf{x} \leq \mathbf{b}\} \geq 1-\varepsilon.
\end{aligned} \qquad \text{(A-54)}
$$

Also,

$$
\underbrace{\mathbf{A}(\zeta)\mathbf{x}+\mathbf{b}(\zeta)}_{\alpha(\mathbf{x})+\beta(\mathbf{x})\in\mathbf{Q}\ \forall\zeta\in\mathcal{Z}}\ , \qquad \text{(A-55)}
$$

where $\mathbf{A}(\zeta) \in \mathcal{R}^{k\times n}$ and $\mathbf{b}(\zeta) \in \mathcal{R}^k$ are affine in ζ. Thus, $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are affine in $\mathbf{x}$ as well.

Consider the perturbation set $\mathcal{Z}$ as a side-wise set $\mathcal{Z} = \mathcal{Z}^{\mathsf{left}} \times \mathcal{Z}^{\mathsf{right}}$, i.e., a product of two sets $\mathcal{Z}^{\mathsf{left}}$ and $\mathcal{Z}^{\mathsf{right}}$, and corresponding to two perturbation vectors $\mathbf{u} \in \mathcal{Z}^{\mathsf{left}}$ and $\mathbf{v} \in \mathcal{Z}^{\mathsf{right}}$, with the left side of conic inequality constraint depending on $\mathbf{u}$, and the right side of the conic inequality constraint depending on $\mathbf{v}$. Therefore, we have:

$$
\|\underbrace{\mathbf{A}(\mathbf{u})\mathbf{x}+\mathbf{b}(\mathbf{u})}_{\equiv\alpha(\mathbf{x})\mathbf{u}+\beta(\mathbf{x})}\| \leq \underbrace{\mathbf{c}^T(\mathbf{v})\mathbf{x}+d(\mathbf{v})}_{\equiv\sigma(\mathbf{x})\mathbf{v}+\delta(\mathbf{x})} \ \ \forall(\mathbf{u}\in\mathcal{Z}^{\mathsf{left}}, \mathbf{v}\in\mathcal{Z}^{\mathsf{right}}). \qquad \text{(A-56)}
$$

Also, we have the norm-bounded perturbation left-side set as

$$
\mathcal{Z}^{\mathsf{left}} = \{\mathbf{u}\in\mathcal{R}^{p\times q} : \|\mathbf{u}\| \leq 1\}. \qquad \text{(A-57)}
$$

B.21 Bisection Methods

Given a continuous function $f(x)$ on the interval $[a, b]$, and $f(a)f(b) < 0$, bisection methods can be used to find a root solution of this function, i.e.,

$$
f(x) = 0. \qquad \text{(A-58)}
$$

Bisection method, also called binary search method, is based on intermediate value theory, which assumes that there exists a unique t such that

$$f(\hat{t}) = 0, \exists t \in [a, b]. \tag{A-59}$$

A bisection method is presented here as Algorithm 1 to find the unique root of the function $f(t)$.

Algorithm 3 Bisection method

1: Set $t_1 = a$, $t_2 = b$;
2: Let $t = \frac{t_1+t_2}{2}$;
3: Calculate $f(t)$;
4: **if** $f(t) = 0$, **then**
5: $\hat{t} = t$;
6: **else**
7: Check the sign of $f(t)$
8: **if** $f(t)$ has the same sign as $f(a)$ **then**
9: $\hat{t} \in [t, b]$, set $t_1 = t$ and go back to Step 2;
10: **else**
11: $\hat{t} \in [a, t]$, set $t_2 = t$ and go back to Step 2;
12: **end if**
13: **end if**

Bibliography

[1] Physical layer security for two way relay communications with friendly jammers. pages 1–6, 2010.

[2] A.Goldsmith and A.Nin. *Wireless Communications*. Cambridge University Press, 2005.

[3] A.Khabbazibasmenj, F.Roemer, S.A.Vorobyov, and M. Haardt. Sum-rate maximization in two-way af mimo relaying: Polynomial time solutions to a class of dc programming problems. *IEEE Trans. Signal Proc.*, 60(10):5478–5493, 2012.

[4] R. Alshwede, N. Cai, S. Li, and R. W. Yeung. Network information flow: Single source. *IEEE Transactions on Information Theory*, 46(4):1204–1216, 2000.

[5] IEEE Standards Association. Ieee standard for local and metropolitan area networks: Overview and architecture. 2002.

[6] IEEE Standards Association. The first ieee workshop on wireless lans: Preface. March 2008.

[7] IEEE Standards Association. Ieee 802.15 wpan task group 6 (tg6) body area networks. July 2011.

[8] Ahsan Aziz, Meng Zeng, Jianwei Zhou, Costas N. Georghiades, and Shuguang Cui. Robust beamforming with channel uncertainty for two-way relay networks. *ICC*, pages 3632–3636, 2012.

[9] A. Ben-Tal, L. El Ghaoui, and A.S. Nemirovski. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press, October 2009.

[10] Mats Bengtsson and Bjorn Ottersten. Optimal downlink beamforming using semidefinite optimization. *37th Annual Allerton Conference on Communication, Control and Computing*, 1999.

[11] A. BenTal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[12] Aggelos Bletsas, Hyundong Shin, and Moe Z. Win. Outage optimality of opportunistic amplify-and-forward relaying. *IEEE Communications Letters*, 11(3):261–263, 2007.

[13] B.Nazer and M.Gastpar. Computing over multiple-access channels with connections to wireless network coding. *Proceedings of the 2006 International Symposium on Information Theory (ISIT 2006)*, July 2006.

[14] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge Uni. Press, 2004.

[15] Batu K. Chalise and Andreas Czylwik. Robust uplink beamforming based upon minimum outage probability criterion. pages 3974–3978. IEEE, 2004.

[16] Deqiang Chen, Kambiz Azarian, and J. Nicholas Laneman. A case for amplify-forward relaying in the block-fading multiple-access channel. *IEEE Transactions on Information Theory*, 54(8):3728–3733, 2008.

[17] J. Chen, R. Zhang, L. Song, Z. Han, and B. Jiao. Joint relay and jammer selection for secure two-way relay networks. *IEEE Int. Conf. Commun.*, Jun. 2011, Kyoto, Japan.

[18] Jingchao Chen, Rongqing Zhang, Lingyang Song, Zhu Han, and Bingli Jiao. Joint relay and jammer selection for secure two-way relay networks. *IEEE Transactions on Information Forensics and Security*, 7(1):310–320, 2012.

[19] Lei Chen, Kai-Kit Wong, Haixia Chen, Ju Liu, and Gan Zheng. Optimizing transmitter-receiver collaborative-relay beamforming with perfect csi. *IEEE Communications Letters*, 15(3):314–316, 2011.

[20] C.Luo, C.Xing, Z.Fei, S.Ma, and J.Kuang. Distributed filter-and-forward beamforming for two-way relaying networks under channel uncertainties. *VTC Spring, IEEE 75th*, May 2012.

[21] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience, New York, NY, USA, 1991.

[22] I. Csiszar and J. Korner. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory*, 24(3):339–348, 1978.

[23] Hongyu Cui, Rongqing Zhang, Lingyang Song, and Bingli Jiao. Performance analysis of bidirectional relay selection with imperfect channel state information. 2013.

[24] Tao Cui and Jörg Kliewer. Memoryless relay strategies for two-way relay channels: Performance analysis and optimization. In *ICC*, pages 1139–1143. IEEE, 2008.

[25] C.Xing, Y.Wu S.Ma, and T.Ng. Robust beamforming for amplify-and-forward mimo relay systems based on quadratic matrix programming. *ICASSP*, pages 3250–3253, 2010.

[26] P. Elias, A. Feinstein, and C. Shannon. A note on the maximum flow through a network. *IEEE Transactions on Information Theory*, 2(4):117–119, 1956.

[27] Siavash Fazeli Dehkordy, Saeed Gazor, and Shahram Shahbazpanahi. Distributed peer-to-peer multiplexing using ad hoc relay networks. IEEE, 2008.

[28] Shengli Fu, Tao Zhang, and Michael Colef. Secrecy in two-way relay systems. IEEE, Dec. 2010.

[29] Kramer G, Gastpar. M, and Gupta.P. Cooperative strategies and capacity theorems for relay networks. *IEEE Transactions on Information Theory*, 51(9):3037–3063, 2005.

[30] E.A. Gharavol and E.Larsson. Robust joint optimization of mimo two-way relay channels with imperfect csi. *IEEE Proc. of Allerton CCC*, 2011.

[31] E.A. Gharavol and E.Larsson. Robust joint optimization of non-regenerative mimo relay channels with imperfect cs. *Signals, Systems and Computers (ASILOMAR)*, 2011.

[32] K. S. Gomadam and S. A. Jafar. Optimal relay functionality for snr maximization in memoryless relay networks. *IEEE J. Sel. Areas Commun.*, 25(2):390–401, 2007.

[33] P. K. Gopala, L. Lai, and H. E. Gamal. On the secrecy capacity of fading channels. *IEEE Trans. Inform. Theory*, 54(10):4687–4698, 2008.

[34] G.Zheng, K.Wong, A.Paulraj, and B. Ottersten. Robust collaborative-relay beamforming. *IEEE Trans. on Signal Proc.*, 57(8), Aug. 2009.

[35] I. Hammerstrom, M. Kuhn, C. Esli, J. Zhao, A. Wittneben, and G. Bauch. Mimo two-way relaying with transmit csi at the relay. pages 1–5, 2007.

[36] V. Havary-Nassab, S. Shahbazpanahi, and A. Grami. Optimal distributed beamforming for two-way relay networks. *IEEE Trans. Sig. Proc.*, 58(3):1238–1250, 2010.

[37] Xiang He and Aylin Yener. The role of feedback in two-way secure communications. *CoRR*, 2009.

[38] Tracey Ho, Muriel Medard, Jun Shi, Michelle Effros, and David R. Karger. On randomized network coding. *In Proceedings of 41st Annual Allerton Conference on Communication, Control, and Computing*, 2003.

[39] Roger A. Horn and Charles R. Johson. *Matrix Analysis*. Cambridge University Press, 1990.

[40] Huang J and A. Lee Swindlehurst. Robust secure transmission in miso channels based on worst-case optimization. *CoRR*, 2011.

[41] Mo J, Liu M, Xia B, and May X. Secure beamforming for mimo two-way transmission with an untrusted relay. *IEEE WCNC*, 2013.

[42] Yindi Jing and Hamid Jafarkhani. Network beamforming using relays with perfect channel information. *IEEE Transactions on Information Theory*, 55(6):2499–2517, 2009.

[43] J.Zhang and M.Gursoy. Collaborative relay beamforming for secrecy. *ICC*, 2010.

[44] I. Karasalo. Estimating the covariance matrix by signal subspace averaging. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 34(1):8–12, 1986.

[45] A. Khisti and G. Wornell. Secure transmission with multiple antennas: The misome wiretap channel. *IEEE Trans. on Information Theory*, 56(7):3088 – 3104, 2010.

[46] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar. On the gaussian mimo wiretap channel. *2007 IEEE International Symposium on Information Theory*, pages 2471–2475, 2007.

[47] S. J. Kim, P. Mitran, C. John, R. Ghanadan, and V. Tarokh. Coded bidirectional relaying in combat scenarios. pages 1–7, 2007.

[48] Johannes Kron, Daniel Persson, Mikael Skoglund, and Erik G. Larsson. Closed-form sum-mse minimization for the two-user gaussian mimo broadcast channel. *IEEE Communications Letters*, 15(9):950–952, 2011.

[49] J. Nicholas Laneman, David N. C. Tse, and Gregory W. Wornell. Cooperative diversity in wireless networks: efficient protocols and outage behavior. *IEEE Trans. Inform. Theory*, 50:3062–3080, 2004.

[50] P. Larsson, N. Johansson, and K. E. Sunell. Coded bidirectional relaying. *IEEE Veh. Technol. Conf. (VTC 2006 Spring)*, 2:851–855, 2006.

[51] S. K. Leung-Yan-Cheong and M. E. Hellman. The gaussian wiretap channel. *IEEE Trans. on Information Theory*, 24(4):451–456, 1978.

[52] Y. Liang and H. V. Poor. Generalized multiple access channels with confidential messages. *Proceedings of IEEE International Symposium on Information Theory*, 2006.

[53] Y. Liang, H. V. Poor, and S. Shamai (Shitz). Secure communication over fading channels. *IEEE Transactions on Information Theory*, 54(6), Jun.2008.

[54] Ruoheng Liu and H Vincent Poor. Secrecy capacity region of a multi-antenna gaussian broadcast channel with confidential messages. *IEEE Transactions on Information Theory*, abs/0709.4(3):1235–1249, 2007.

[55] Ruoheng Liu and Wade Trappe. *Securing Wireless Communications at the Physical Layer*. springer, 2010.

[56] M.Tao and R. Wang. Robust relay beamforming for two-way relay networks. *IEEE Comm. Letters*, 16(7):1052–1055, 2012.

[57] A. Mukherjee and A. Lee Swindlehurst. Securing multiantenna two-way relay channels with analog network coding against eavesdroppers. *Sig. Proc. Advances Wireless Commun.*, pages 1–5, 2010.

[58] Amitav Mukherjee and A. Lee Swindlehurst. Utility of beamforming strategies for secrecy in multiuser mimo wiretap channels. *Conference on Communication, control, and computing*, 1134 - 1141, 2009.

[59] M.Zeng, R.Zhang, and S.Cu. On design of distributed beamforming for two-way relay networks. *IEEE Transactions on Signal Processing*, 59(9):2284 – 2295, May 2011.

[60] F. Oggier and B. Hassibi. The secrecy capacity of the mimo wiretap channe. *IEEE Int. Symp. Information Theory*, 2008.

[61] Agisilaos Papadogiannis, Alister G. Burr, and Meixia Tao. On the maximum achievable sum-rate of interfering two-way relay channels. *IEEE Communications Letters*, 16(1):72–75, 2012.

[62] P. Parada and R. Blahut. Secrecy capacity of simo and slow fading channels. *IEEE International Symposium on Information Theory*, 2005.

[63] C S Patel, G L Stuber, and T G Pratt. Statistical properties of amplify and forward relay fading channels. *IEEE Transactions on Vehicular Technology*, 55(1):1–9, 2006.

[64] A.J. Pierrot and M.Bloch. Strongly secure communications over the two-way wiretap channel. *IEEE Transactions on Information Forensics and Security*, 6(3):595–605, 2011.

[65] B. Rankov and A. Wittneben. Achievable rate regions for the two way relay channel. *Information Theory, 2006 IEEE International Symposium on*, 2006.

[66] R.H. Roy and B. Ottersten. Spatial division multiple access wireless communication systems. 1991. US Patent 5,515,378.

[67] C. Schnurr, S. Stanczak, and T. J. Oechtering. Coding theorems for the restricted half-duplex two-way relay channel with joint decoding. *IEEE Int. Symp. Inf. Theory*, pages 2688–2692, Jul. 2008.

[68] Andrew Sendonaris, Elza Erkip, and Behnaam Aazhang. User cooperation diversity - part i: System description. *IEEE Trans. Commun*, 51:1927–1938, 2003.

[69] Andrew Sendonaris, Elza Erkip, and Behnaam Aazhang. User cooperation diversity - part ii: Implementation aspects and performance analysis. *IEEE Transactions on Communications*, 51:1939–1948, 2003.

[70] S. Shahbazpanahi and M. Dong. A semi-closed-form solution to the snr balancing problem of two-way relay network beamforming. *IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 2010.

[71] C. Shannon. Communication theory of secrecy systems. October 1949.

[72] Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, July, October 1948.

[73] Takayuki Shimizu, Hisato Iwai, and Hideichi Sasaoka. Physical-layer secret key agreement in two-way wireless relaying systems. *IEEE Transactions on Information Forensics and Security*, 6(3-1):650–660, 2011.

[74] S.Katti, S. Gollakota, and D.Katabi. Embracing wireless interference: Analogue network coding. *ACM SIGCOMM*, pages 397–408, 2007.

[75] S.Li, Raymond W. Yeung, and Ning Cai. Linear network coding. *IEEE Transactions on Information Theory*, 49:371–381, 2003.

[76] Sturm. Using sedumi 1.02, a matlab toolbox or optimization over symmetric cones. *Optimization Methods and Software*, 11, 1999.

[77] J. F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *SEEM2001-01*, 2001.

[78] S.Zhang and S.-C. Liew. The capacity of two way relay channel. *4th International Conference on Access Networks AccessNets*, pages 219–231, 2008.

[79] Ender Tekin and Aylin Yener. The general gaussian multiple-access and two-way wiretap channels: Achievable rates and cooperative jamming. *IEEE Transactions on Information Theory*, 54(6):2735–2751, 2008.

[80] P. Ubaidulla and A. Chockalingam. Robust distributed beamforming for wireless relay networks. *Personal, Indoor and Mobile Radio Communications, IEEE 20th International Symposium on*, pages 2345 – 2349, Sept. 2009.

[81] Barry van Veen and Kevin M. Buckley. Beamforming: A versatile approach to spatial filtering. *IEEE ASSP Magazine*, 5:4–24, April 1988.

[82] R. Vaze and R. W. Heath. Capacity scaling for mimo two-way relaying. *IEEE Int. Sym. Info. Theory*, 2007.

[83] Fanggang Wang, Xiaojun Yuan, Soung Chang Liew, and Dongning Guo. Wireless mimo switching: Weighted sum mean square error and sum rate optimization. *IEEE Transactions on Information Theory*, 59(9):5297–5312, 2013.

[84] Huiming Wang, Qinye Yin, and Xiang-Gen Xia. Distributed beamforming for physical-layer security of two-way relay networks. *IEEE Transactions on Signal Processing*, 60(7):3532–3545, 2012.

[85] Wenjin Wang, Shi Jin, and Fu-Chun Zheng. Maximin snr beamforming strategies for two-way relay channels. *IEEE COMMUNICATIONS LETTERS*, 16(7):1006–1009, 2012.

[86] W.Nam, S. Chung, and Yong H. Lee. Capacity of the gaussian two-way relay channel to within 1/2 bit. *IEEE Trans. Inf. Theory*, 56(11):5488 – 5494, 2010.

[87] A. Wyner. The wiretap channel. *Bell. Sys. Tech. J.*, 54(8):1355–1387, Jan. 1975.

[88] Y. Ye and S. Zhang. New results on quadratic minimization. *SIAM J. Optim.*, 14(1):245–267, 2003.

[89] W. Yu and J. Cioffi. Sum capacity of gaussian vector broadcast channels. *IEEE Trans. Info. Theory*, 50(9):1875–1892, Sept. 2004.

[90] C. Yuen, W. H. Chin, Y. L. Guan, W. Chen, and T. Tee. Bidirectional multi-antenna relay communications with wireless network coding. *IEEE Veh. Technol. Conf. (VTC-Spring 2008)*, pages 1385–1388, May 2008.

[91] J. Zhang and M. C. Gursoy. Collaborative relay beamforming for secure broadcasting. *Wireless Commun. Net. Conf.*, pages 1–6, Apr. 2010.

[92] Jianshu Zhang, Florian Roemer, Martin Haardt, Arash Khabbazibasmenj, and Sergiy A. Vorobyov. Sum rate maximization for multi-pair two-way relaying with single-antenna amplify and forward relays. In *ICASSP*, pages 2477–2480. IEEE, 2012.

[93] R. Zhang, Y. Liang, C. Chai, and S. Cui. Optimal beamforming for two-way multi-antenna relay channel with analogue network coding. *IEEE J. Select. Areas Commun*, 27(5):669–712, 2009.

[94] Shengli Zhang. Hot topic: physical-layer network coding. *ACM Mobicom*, pages 358–365, 2006.

[95] G. Zheng, L. C. Choo, and K. K. Wong. Optimal cooperative jamming to enhance physical layer security using relays. *IEEE Trans. Sig. Proc.*, 59(3):1317–1322, Mar. 2011.

[96] Jun Zou, H.Luo, M.Tao, and R.Wang. Joint source and relay optimization for non-regenerative mimo two-way relay systems with imperfect csi. *IEEE Trans. Wireless Comm.*, 11(9):3305–3315, 2012.